

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2005, 09:06:37 ; Search time 44 Seconds
(without alignments)
1812.811 Million cell updates/sec

Title: US-10-696-639-39
Perfect score: 4369
Sequence: 1 MGLPRGPLASLLQLQVCWLQ.....NEWGSRFKKLADMYGGGEDD 829
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4369	100.0	829	1 IJHUCP	cadherin 3 precurs
2	3555.5	81.4	822	1 IJMSCP	p-cadherin precurs
3	2573.5	58.9	732	1 IJHCEB	B-cadherin precurs
4	2420	55.4	882	1 IJHUCE	cadherin 1 precurs
5	2388	54.7	884	1 IJM5CE	uvomorulin - mouse
6	2379	54.5	884	2 S34438	EP-cadherin precurs
7	2375	54.4	895	1 IJXLCP	cadherin - African
8	2371	54.3	905	2 S43064	p-cadherin - bovin
9	2323	53.2	491	1 IJBQCP	E-cadherin precurs
10	2274	52.0	887	1 IJHCHL	cadherin - African
11	2192.5	50.2	871	2 S47518	cadherin 2 precurs
12	1707	39.1	906	1 IJHUCN	N-cadherin precurs
13	1704.5	39.0	912	1 IJHCHN	N-cadherin precurs
14	1697.5	38.9	906	1 IJM5CN	R-cadherin precurs
15	1694	38.8	913	1 IJHCHR	N-cadherin precurs
16	1691	38.7	913	1 IJBQCN	R-cadherin precurs
17	1655	37.9	913	1 A47543	cadherin 4 precurs
18	1621.5	37.1	916	2 C38992	N-cadherin precurs
19	1621	37.1	783	2 I50116	N-cadherin 2 precu
20	1616.5	37.0	906	1 IJXL2	N-cadherin 1 precu
21	1609	36.8	905	1 IJXL1	cadherin-15 precu
22	1321.5	30.2	814	2 G02878	M-cadherin - mouse
23	1302.5	29.8	730	1 IJM5CM	cadherin 13 precu
24	1021.5	23.4	713	2 B38992	cadherin-7 - chick
25	1007.5	23.1	785	2 I50180	cadherin-6B - chic
26	990	22.7	790	2 I50178	cadherin-6 - human
27	982	22.5	790	2 I37016	cadherin-14 - huma
28	976.5	22.4	790	2 G02678	T-cadherin precurs
29	976	22.3	712	1 IJM5CT	

RESULT 1
IJHUCP

N;Alternate names: P-cadherin; placental cadherin
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C;Accession: A33659
R;Shimoyama, Y.; Yoshida, T.; Terada, M.; Shimamoto, Y.; Abe, O.; Hirohashi, S.
J. Cell Biol. 109, 1787-1794, 1999
A;Title: Molecular cloning of a human Ca(2+)-dependent cell-cell adhesion molecule homol
A;Reference number: A33659; MUID:90009051; PMID:2793940
A;Accession: A33659
A;Molecule type: mRNA
A;Residues: 1-829 <SHI>
A;Cross-references: UNIPROT:P22223; UNIPARC:UPI000004C15E; CB:X63629; NID:935322; PIDN:C
C;Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought t
C;Genetics:
A;Gene: GDB:CDH3

ALIGNMENTS

30	976	22.3	717	2	I51206	T-cadherin 2 - chi
31	974	22.3	789	2	I52701	k-cadherin - rat
32	966.5	22.1	796	2	A38992	cadherin 11 precur
33	963	22.0	863	1	IJBODC	desmocollin 2a pre
34	963	22.0	896	2	A55363	desmocollin, type
35	953.5	21.8	796	2	I48277	cadherin-11 - mous
36	951.5	21.8	796	2	A53584	OB-cadherin precur
37	948.5	21.7	901	1	IJHUDA	desmocollin 3a pre
38	941.5	21.5	796	2	I49556	cadherin-11 - mous
39	941.5	21.5	809	1	IJBODD	desmocollin 2b pre
40	940.5	21.5	896	2	I45858	desmocollin - bovi
41	939.5	21.5	847	1	IJHUBD	desmocollin 3b pre
42	937	21.4	826	2	B55363	desmocollin, type
43	926.5	21.2	793	2	D38992	cadherin 8 - huma
44	924.5	21.2	794	2	I59372	cadherin 12 - huma
45	910	20.8	839	1	IJBODF	desmocollin 1b pre

Query Match 100.0%; Score 4369; DB 1; Length 829;
Best Local Similarity 100.0%; Pred. No. 1.6e-270; Indels 0; Gaps 0;
Matches 829; Conservative 0; Mismatches 0;

Qy	1	MGLPRGPLASLLQLQVCWLQCAASEPCRAVFREAEVTLAAGAEQEQALGVFMGCPG	60
Db	1	MGLPRGPLASLLQLQVCWLQCAASEPCRAVFREAEVTLAAGAEQEQALGVFMGCPG	60
Qy	61	QEPALSTDNDDFTVRNGETVQERRSLKERNPLKIPFSGILRRHKRDWVAPISVPENG	120
Db	61	QEPALSTDNDDFTVRNGETVQERRSLKERNPLKIPFSGILRRHKRDWVAPISVPENG	120
Qy	121	KGPPORLNQKSKNDKDTKIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDREBIAK	180
Db	121	KGPPORLNQKSKNDKDTKIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDREBIAK	180

```
181 YELFGHAVSENGASVEDPMNISIIVTDNDHKPKFTQDTERGSLVGLVPGTSMQVMTAT 240
181 YELFGHAVSENGASVEDPMNISIIVTDNDHKPKFTQDTERGSLVGLVPGTSMQVMTAT 240
241 DEDDAIYTYNGVAVSISHSQEPKDPHDLMTIHRSTGTISVISGGLDREKVPEYTLTIOA 300
241 DEDDAIYTYNGVAVSISHSQEPKDPHDLMTIHRSTGTISVISGGLDREKVPEYTLTIOA 300
301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHEVORLTVTDLDPNSP 360
301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHEVORLTVTDLDPNSP 360
361 AWRATYLMGDDGDHFTIITHPESNOGILTTTRKGLDFAKNQHTLYVETNAPFVLKL 420
361 AWRATYLMGDDGDHFTIITHPESNOGILTTTRKGLDFAKNQHTLYVETNAPFVLKL 420
421 PTSTATIVVHVEDVNEAPVFPSPKVEVEGIPTEGFCVYTAEDPDKENQKISYRILR 480
421 PTSTATIVVHVEDVNEAPVFPSPKVEVEGIPTEGFCVYTAEDPDKENQKISYRILR 480
481 DPAGWLAMDPSGQVTAAGTLDREDEQFVRNNIYEVNVLAMDNGSPPTTGTGTLTLLID 540
481 DPAGWLAMDPSGQVTAAGTLDREDEQFVRNNIYEVNVLAMDNGSPPTTGTGTLTLLID 540
541 VNDHGPPEPROITICNOSPVRHVLNITDKOLSPHTSPFOAQLTDDSDIYMTAEVNEEGD 600
541 VNDHGPPEPROITICNOSPVRHVLNITDKOLSPHTSPFOAQLTDDSDIYMTAEVNEEGD 600
601 TVVLSLKKFLKQDYDVHLSLSHGNKEQLTVIRATVCDCHGVETCPGPKWGGFILPVL 660
601 TVVLSLKKFLKQDYDVHLSLSHGNKEQLTVIRATVCDCHGVETCPGPKWGGFILPVL 660
661 GAVLALLFLLVLLVLLVLRKKRKIKEPLLLPEDDTRDNVFFYGGEGGEDQDYDITQHR 720
661 GAVLALLFLLVLLVLLVLRKKRKIKEPLLLPEDDTRDNVFFYGGEGGEDQDYDITQHR 720
721 GLEARPEVLNVDNAPFTIIPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
721 GLEARPEVLNVDNAPFTIIPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
781 DYEGSGSDAASLSLTSASDQDQDYDYLNEWGSRRFKKLADMYGGGEDD 829
781 DYEGSGSDAASLSLTSASDQDQDYDYLNEWGSRRFKKLADMYGGGEDD 829

RESULT 2
IJMSCP
P-cadherin precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C;Accession: S03163; S34458
R;Nose, A.; Nagafuchi, A.; Takeichi, M.
EMBO J. 6, 3655-3661, 1987
A;Title: Isolation of placental cadherin cDNA: identification of a novel gene family of
A;Reference number: S03163; MUID:88111554; PMID:3428270
A;Accession: S03163
A;Molecule type: mRNA
A;Residues: 1-822 <NOS>
A;Cross-references: UNIPROT:P10287; UNIPARC:UPI000017435B; EMBL:X06340
R;Faraldo, M.L.M.; Cano, A.
J. Mol. Biol. 231, 935-941, 1993
A;Title: The 5' flanking sequences of the mouse P-cadherin gene. Homologies to 5' sequen
A;Reference number: S34458; MUID:93294853; PMID:8515462
A;Accession: S34458
A;Molecule type: DNA
A;Residues: 1-55 <PAR>
A;Cross-references: UNIPARC:UPI000016CCF0; EMBL:X68057
C;Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought t
C;Genetics:
A;Intons: 16/3
C;Superfamily: cadherin; cadherin repeat homology
C;Keywords: calcium binding; cell adhesion; duplication; glycoprotein; placenta; transme
```

```
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-99/Domain: propeptide #status predicted <PRO>
F;100-822/Product: P-cadherin #status predicted <MAT>
F;100-645/Domain: extracellular #status predicted <EXT>
F;102-207/Domain: cadherin repeat homology <CR1>
F;210-320/Domain: cadherin repeat homology <CR2>
F;323-432/Domain: cadherin repeat homology <CR3>
F;433-540/Domain: cadherin repeat homology <CR4>
F;541-645/Domain: cadherin repeat homology <CR5>
F;646-670/Domain: transmembrane #status predicted <TMW>
F;671-822/Domain: intracellular #status predicted <INT>
F;778-793/Region: serine-rich
F;192,558/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 81.4%; Score 3555.5; DB 1; Length 822;
Best Local Similarity 81.6%; Pred. No. 1.2e-218; Indels 13; Gaps 5;
Matches 679; Conservative 57; Mismatches 83;

QY 1 MGLPRGPLA-SLLLLQVLCWLAASBP CRAVF-REAEVTLEAGABEQEQALGKVP MGC 58
Db 1 MELLSGPHAFLLLLQVLCWLSVVSFYPAGFIGEAGVTLEVEGTDLEPSQVLGKVALAG 60

QY 59 PQQEPALFSTDDNDFTRNGETVQERRSLKERNPLKIFSKRILRRHKRDVWVAPISVPE 118
Db 61 QGMHHA---DNGDIIIMLTRGTQVGKDAHMS-----PPTRILRRRKREWVMPPIFVPE 110

QY 119 NGKGPORLNOLKSNKRDTKIFYSTIGPGADSPPEGVFAVEKETGWLILNKLPLDREEI 178
Db 111 NGKGPORLNOLKSNKRDGKIFYSTIGPGADSPPEGVFTI EKESGWLILHLPDLREKI 170

QY 179 AKYELFGHAVSENGASVEDPMNISIIVTDNDHKPKFTQDTERGSLVGLVPGTSMQVMT 238
Db 171 VKYELFGHAVSENGASVESEPMNISIIVTDNDNKPFTQDTERGSLVGLVPGTSMQVMT 230

QY 239 ATDEDDAIYTYNGVAVSISHSQEPKDPHDLMTIHRSTGTISVISGGLDREKVPEYTLTI 298
Db 231 ATDEDDAVNTYNGVAVSISHSQEPKDPHDLMTIHRSTGTISVISGGLDREKVPEYTLTV 290

QY 299 QATMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHEVORLTVTDLDPN 358
Db 291 QATMDGSGSTTTAVAVVEILDANDNAPMFEPOKYEAWPENEVGHVORLTVTDLDPN 350

QY 359 SPAWRATYLMGDDGDHFTIITHPESNOGILTTTRKGLDFAKNQHTLYVETNAPFVL 418
Db 351 WPAWRATYHIVGGDDGDHFTIITHPETNQGVLLTKKGLDFAEQDQHTLYVETNAPFAV 410

QY 419 KLPTSTATIVVHVEDVNEAPVFPSPKVEVEGIPTEGFCVYTAEDPDKENQKISYRI 478
Db 411 KLPTATATVYVHVKVDNEAPVFPSPKVEVEGIPTEGFCVYTAEDPDKENQKISYTI 470

QY 479 LRDPAGWLAMDPSGQVTAAGTLDREDEQFVRNNIYEVNVLAMDNGSPPTTGTGTLTLL 538
Db 471 SRDPANWLAVDPDSGQITAAAGILDREDEQFVRNNIYEVNVLAMDNGSPPTTGTGTLTLL 530

QY 539 IDVNDHGPPEPROITICNOSPVRHVLNITDKOLSPHTSPFOAQLTDDSDIYMTAEVNEE 598
Db 531 TDINDHGPPEPROITICNOSPVPQVNLNITDKOLSPNSPFOAQLTHDSDIYMAVSEK 590

QY 599 GDTVVLGKFLKQDYDVHLSLSHGNKEQLTVIRATVCDCHGV-ETCGPKWGGFIL 657
Db 591 GDTVALSLKFLKQDYDYLHLSLSDHGNREQLTMIRATVCDCHGVFNDCPRPKWGGFIL 650

QY 658 PVLGAVLALLFLLVLLVLLVLRKKRKIKEPLLLPEDDTRDNVFFYGGEGGEDQDYDITQ 717
Db 651 PILGAVLALLFLLVLLVLLVLRKKRKIKEPLLLPEDDTRDNVFFYGGEGGEDQDYDITQ 710

QY 718 LHRGLEARPEVLNVDNAPFTIIPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTL 777
Db 711 LHRGLEARPEVLNVDNAPFTIIPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDSL 770

QY 778 LVFDYEGSGSDAASLSLTSASDQDQDYDYLNEWGSRRFKKLADMYGGGEDD 829
Db 771 MVFDYEGSGSDAASLSLTSASDQDQDYDYLNEWGSRRFKKLADMYGGGEDD 822
```

RESULT 3

IJCCHB

B-cadherin precursor - chicken (fragment)
N/Alternate names: K-CAM protein
C/Species: Gallus gallus (chicken)
C/Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 09-Jul-2004
C/Accession: A41634; A38715; S16160
R/Sorkin, B.C.; Gallin, W.J.; Edelman, G.M.; Cunningham, B.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 11545-11549, 1991
A/Title: Genes for two calcium-dependent cell adhesion molecules have similar structures
A/Reference number: A41634; MUID:92107987; PMID:1763068
A/Accession: A41634
A/Molecule type: DNA
A/Residues: 1-732 <SOR>
A/Cross-references: UNIPROT:P33145; UNIPARC:UPI0000136DBF; GB:M81894; NID:g212226; PIDN:
R/Napolitano, E.W.; Venstrom, K.; Wheeler, E.F.; Reichardt, L.F.
J. Cell Biol. 113, 893-905, 1991
A/Title: Molecular cloning and characterization of B-cadherin, a novel chick cadherin.
A/Reference number: A38715; MUID:91225083; PMID:2026653
A/Accession: A38715
A/Molecule type: mRNA
A/Residues: 7-413, 'V', 415-732 <NAP>
A/Cross-references: UNIPARC:UPI00001712B4; GB:X58518; NID:g63113; PIDN:CAA41408.1; PID:9
C/Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought b
C/Genetics:
A/Gene: K-CAM
A/Introns: 29/3; 81/3; 130/1; 188/3; 231/3; 293/3; 375/2; 423/1; 498/1; 571/1; 614/3; 66
C/Superfamily: cadherin; cadherin repeat homology
C/Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro
F/1-6/Domain: propeptide (fragment) #status predicted <PRO>
F/6-554/Domain: extracellular #status predicted <EXT>
F/7-732/Product: B-cadherin #status predicted <MAT>
F/9-114/Domain: cadherin repeat homology <CR1>
F/84-89/Region: cadherin binding #status predicted
F/117-227/Domain: cadherin repeat homology <CR2>
F/230-339/Domain: cadherin repeat homology <CR3>
F/340-447/Domain: cadherin repeat homology <CR4>
F/448-552/Domain: cadherin repeat homology <CR5>
F/555-580/Domain: transmembrane #status predicted <TM>
F/581-732/Domain: intracellular #status predicted <INT>
F/689-702/Region: serine-rich
F/137,410/Binding site: carboxydrate (Asn) (covalent) #status predicted

Query Match 58.9%; Score 2573.5; DB 1; Length 732;
Best Local Similarity 66.2%; Pred. No. 4e-156;
Matches 485; Conservative 103; Mismatches 138; Indels 7; Gaps 5;

QY 102 LRRHKRDWVAPISVSPENGKGPFPQRLNQLKSKNRDRTKIFYISITGPGADSPPEGVFAVE 161
DB 1 LRQKRDWVPPKVPENEGPFPKNLVQIKSNRDREAKIFYISITGGADAPPEGIFTIE 60
QY 162 KETGWLKLPDREIAKYELFGHAVSENGASVEDPMNISIIVTQNDHKPKFTQDTPR 221
DB 61 KETGWNKVTQPLDREHINKYLYSHAVSENGKVPSEPMIIVTQNDKPKQFTQEVER 120
QY 222 GSVLEGLPCTSNVQVATDEDAIYTYNGVAVSYHSQEPKDPHDLMTIHRSTGTSV 281
DB 121 GSVPEGALPGTSMVRNATDADDVETNGVAVSYLSQEPHPHMFVNATGTLV 180
QY 282 ISSGLDREKPEYTLTIQATDMGDGTTTAVAVVELDANDNAPMFDPOKYEAHPENA 341
DB 181 IASGLDRERVRETLTQWADLDGQGLTTTALAVIEITVDNDNAPFDPKTYEAAVENE 240
QY 342 VGHEVQRVLTDLADPNAPWATYLLIMGGDDGHDFTTTHPESNQGLTRKGLDFEAK 401
DB 241 AELEVARLATDDEHPAPWAVSYVRNGEGAFITTTDPASNEGLRTAKGLDYEAK 300
QY 402 NQHTLYVEVTNEAPFVKLPSTATIVVHVEDYNEAPVFPSPKVEVQEGITGEPCV 461
DB 301 RQFVLHVAVVNEAPFAIKLPATATATVMVSDYNEAPVFPDPLRLAQVPEDVPLGQPLAS 360

QY 462 YTAEDPDK-ENOKISYRILRDPAGWLANDPDSGGQVTAAGTLDREDEQFVRNNIYVMVLA 520
DB 361 YTAQDPDRAQQORIKYVMGSDPDAGWLAHPENGLITAREQLDRE-SPETKNSTYMAVLA 419
QY 521 WDNGSPPTTGTGTLTLLIDVNDHGPVPEPRQITTCNOSPVRHVLNITDKDLSPHTSPEQ 580
DB 420 VVDGLPPATGTGTLTLLIDVNDHGPVPEPRDIVCNKSPVQVLTITDRLPPNTGPR 479
QY 581 AQLTDDSDIYTAAYNEEGDVTVLSSKKFLKQDQTYDHLHSLSDHGNKEQLTVIRATVDC 640
DB 480 AELSHGSGDSMAVEGVNGGDTVALMLTEPLEQLNYSVYLRLFDROGKQDQTVIRAQCDC 539
QY 641 HGHVETC---POPWKG-GFILPVGLAVLALLLLELLVLLLVKRRKIKKEPILLPEDDTRD 696
DB 540 QGRVESCAQKPRVDTGVPIVLAVLGAVLALLLLELLVLLLVKRRKIKKEPILLPEDDTRD 599
QY 697 NVFYVYEGEGGEDDYDITQLHRGLEARPEVVLRNDVAPTITPTPMYPRPANPDEIGN 756
DB 600 NIFYVYEGEGGEDDYDUSLHRGLDARPEVI-RNDVAPPLMAAPQYRPRPANPDEIGN 658
QY 757 FIENLKAANTDPTAPPYDTLLVFDYEGSGDAASLSLTSSASDQDQDYDYLNEWGSRF 816
DB 659 FIDENLKAANTDPTAPPYDLSLVFDYEGSGSEATSLSLNSSASDQDQDYDYLNEWGSRF 718
QY 817 KKLADMYGGEDD 829
DB 719 KKLAEYGGGEDE 731

RESULT 4

IJCCHB

cadherin 1 precursor [validated] - human
N/Alternate names: ARC-1; cell CAM 120/80; E-cadherin; epithelial cadherin; L-CAM; uvomor
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text change 09-Jul-2004
C/Accession: S37654; S31430; S05475; S31460; S06716; AS171; JC2230; I52294; I52704; S255
R/Busemackers, M.J.G.; van Bokhoven, A.; Mees, S.G.M.; Kemler, R.; Schalken, J.A.
Mol. Biol. Rep. 17, 123-128, 1993
A/Title: Molecular cloning and characterization of the human E-cadherin cDNA.
A/Reference number: S37654; MUID:93211394; PMID:8459805
A/Accession: S37654
A/Molecule type: mRNA
A/Residues: 1-882 <BUS>
A/Cross-references: UNIPROT:P12830; UNIPARC:UPI0000341BF; EMBL:Z13009; NID:g31072; PIDN:
R/Keller, W.; Warda, A.; Oda, T.; Hirohashi, S.; Kemler, R.; Birchmeier, W.
submitted to the EMBL Data Library, December 1992
A/Description: Sequence of human E-cadherin cDNA.
A/Reference number: S31430
A/Accession: S31430
A/Molecule type: mRNA
A/Residues: 1-542, 'F', 544-882 <LE>
A/Cross-references: UNIPARC:UPI000016A86F; EMBL:Z18923; NID:g31074; PIDN:CAA79356.1; PID
R/Mansouri, A.; Spurr, N.; Goodfellow, P.N.; Kemler, R.
Differentiation 38, 67-71, 1998
A/Title: Characterization and chromosomal localization of the gene encoding the human cel
A/Reference number: S05475; MUID:89031725; PMID:3263290
A/Accession: S05475
A/Molecule type: mRNA
A/Residues: 157-311 <MAN>
A/Cross-references: UNIPARC:UPI0000174350; EMBL:X12790
A/Note: nucleotide sequence is not complete
R/Frixen, U.H.
submitted to the EMBL Data Library, March 1990
A/Reference number: S31460
A/Accession: S31460
A/Molecule type: mRNA
A/Residues: 265-392 <PRI>
A/Cross-references: UNIPARC:UPI000016A583; EMBL:X52279; NID:g28821; PIDN:CAA36522.1; PID
R/Wheelock, M.J.; Buck, C.A.; Bechtol, K.B.; Damsky, C.H.
J. Cell. Biochem. 34, 187-202, 1987
A/Title: Soluble 80-kd fragment of cell-CAM 120/80 disrupts cell-cell adhesion.
A/Reference number: S06716; MUID:87280410; PMID:3611200
A/Accession: S06716

A;Reference number: S03160; MUID:88111553; PMID:3501370
A;Accession: S03160
A;Molecule type: mRNA
A;Residues: 157-884 <RIN>
A;Cross-references: UNIPARC:UPI0000174353; EMBL:X06339
A;Note: part of this sequence, including the amino end of the mature protein, was confirmed by the following authors: Klein-Hitpass, L.; Birchmeier, W.
Proc. Natl. Acad. Sci. U.S.A. 88, 11495-11499, 1991
A;Title: The E-cadherin promoter: Functional analysis of a G-C-rich region and an epithelial enhancer
A;Reference number: I49565; MUID:92107977; PMID:1763063
A;Accession: I49565
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-15 <RES>
A;Cross-references: UNIPARC:UPI000016CC60; GB:M81449; NID:G192325; PIDN:AAA37352.1; PID:FEBS Lett. 352, 318-322, 1994
A;Title: Purification and spectroscopic characterization of a recombinant amino-terminal domain of E-cadherin
A;Reference number: S48735; MUID:95010732; PMID:7925993
A;Accession: S48735
A;Status: preliminary
A;Molecule type: protein
A;Residues: 156-300 <TON>
A;Cross-references: UNIPARC:UPI0000174354
C;Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to be involved in a variety of cellular processes.
C;Genetics:
A;Gene: E-cadherin
C;Superfamily: cadherin; cadherin repeat homology
C;Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane protein
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-156/Domain: propeptide #status predicted <PRO>
F;157-884/Product: E-cadherin, epithelial #status experimental <MAT>
F;157-699/Domain: extracellular #status predicted <EXT>
F;159-264/Domain: cadherin repeat homology <CR1>
F;234-239/Region: cadherin binding #status predicted
F;267-377/Domain: cadherin repeat homology <CR2>
F;380-488/Domain: cadherin repeat homology <CR3>
F;489-597/Domain: cadherin repeat homology <CR4>
F;598-702/Domain: cadherin repeat homology <CR5>
F;702-733/Domain: transmembrane #status predicted <TM>
F;734-884/Domain: intracellular #status predicted <INT>
F;842-855/Region: serine-rich
F;560,639/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 54.6%; Score 2388; DB 1; Length 884;
Best Local Similarity 53.6%; Pred. No. 3.5e-144;
Matches 472; Conservative 123; Mismatches 220; Indels 66; Gaps 13;

QY 9 ASLLLLQV-CWLOCAASEP--CRAVFREAEVTLAAGAEQEPQALGKV-FMGCPOGEPA 64
DB 10 ALLLLLQVSSWL-CQELPESCPGFSSEVTFPPVPERHLERGHVLRVFEFGCTGRPT 68
QY 65 LFSTDNDDFVRNGEIVQERSLK----- 88
DB 69 AFFSEDSRFKATDGTITVTKRHLKHLKLETSFLVRARDSSHRELSTKVTLSKMGHHHRH 128
QY 89 -----ERNP-LKIFPPSKRI-LRRHKRDWVAPISVPENGKGFPPQRLNQLKSNKDRDT 139
DB 129 HHRDPASESNPELLMFVSFVPLGRQRQKDWIPIPCPENEKEGFPKQLVQIKSNDRKET 198
QY 140 KIFYSITGQADSPGCVAFVEKETGMLLNKPLDREIEIAKELFGHAVSENGASVEDPM 199
DB 189 KVFYSITGQADKPPVGVFTIERETGWLKVTQPLDREAIAKYILSHAVSNGEAVEDPM 248
QY 200 NISIVTDNDHKPKFTQDFRGSVLEGLVPGTSMQVATDDEDDALTYNGVAVYSIHS 259
DB 249 EIVITVTDQNDNRPEFTQEFYEGSVAEGAVPGTSVMKVSATDADDVNTYNAIAIYIVS 308
QY 260 QEPKDPHDLMTIHRSTGTTISVTSGLDREKVPETLITIQATDMDGSGSTTAVAVVEIL 319
DB 309 QDELPKHKMFTVNRDGTGLSVITSLGLDRESYFTYTLVQVADLQEGSLSTTAKAVITVK 368
QY 320 DANDNAPMFDPOKYEAHVPENAVGHEVQRLTVTDLDPNSPAWRATYILMGDDGDGHFTI 379

DB 369 DINDNAPFNPSTYQGVPENEVNARIATLTKVTDDAPNTPAWKAVYTVV-NDPOQFVV 427
QY 380 TTHPESNOGILTRKGLDPEAKNQHTLVYVEVNEAPFVLKLTPTSTATIVVHVEDVNEAPV 439
DB 428 VTDPITNDGILKATAGLDFEAKQVILHVRVNEEFEGSLVPSTATVTVDDVNEAPI 487
QY 440 FVPPSKVVEVQGIPTGPEVCVYTAEDPK-BNQKISVRIILDPAGWLAMDPSQGVAV 498
DB 488 FMPAERRVEVPDFFGVGBEITSYAREPDTFMDQKITVIRMRDTANWLEINPETGAIFTR 547
QY 499 GTLDREDEQFVNNIYEVNVLAMDNGSPPTCTGTLLTLIDVNDHGPVPERQITICNQ 558
DB 548 AEMDRDEAHVKNSTVALIITDDGSPATGTGILLVLLVDNNAPIPERNNWQFCOR 607
QY 559 SPVRHVLNITDGLSPHTSPFOQLTDDSDIYWTAEVNEEG-DTVVLSLKELKQDTYDV 617
DB 608 NQPHIITLDPDLPNTSPFTAEILTHGASVNMVITVYNDAAQESLILQPKLEIGEYKI 667
QY 618 HUSLDHGNKEQLTVIRATVCDCHGHVETCPGFWKGF-----ILPGLAVLALLF 668
DB 668 HLKLDNQNKQVTTLDVHVCDCEGTNNC---MKAGIVAAGLQVPAIILGILGILALLI 724
QY 669 LLLVLLLVKRRKIKPELPLLEDTRDNVYVYEGEGGEDDYDITOLHRLGLARPEV 728
DB 725 LILLVLLVRRRTVVKPELPPDDTRDNVYVYDEGEGGEDDPLSLQHLGLDARPEV 784
QY 729 VLNRNDVAPITPTMYRPRPNDPDEIGNFIENLKAANTDPTAPPYDTLLVFDYEGSGD 788
DB 785 T-RNDVAPFLMSVQVPRPNDPDEIGNFIDENLKAADSPTAPPYDLSLLVFDYEGSGSE 843
QY 789 AASLSLSSASDQDQDYDLNWSGRFKKLADMYGGGEDD 829
DB 844 AASLSLSSASDQDQDYDLNWSGRFKKLADMYGGGEDD 884

RESULT 6
34438
uvomorulin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S34438
R;Ringwald, M.; Baribault, H.; Schmidt, C.; Kemler, R.
Nucleic Acids Res. 19, 6533-6539, 1991
A;Title: The structure of the gene coding for the mouse cell adhesion molecule uvomorulin
A;Reference number: S34438; MUID:92093614; PMID:1754391
A;Accession: S34438
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-884 <RIN>
A;Cross-references: UNIPARC:UPI0000020A82; EMBL:X60975
A;Note: the nucleotide sequence was submitted to the EMBL data library, August 1992
C;Genetics:
A;Intons: 16/3; 57/1; 131/3; 179/3; 231/3; 280/1; 338/3; 381/3; 442/3; 524/2; 573/1; 648/1
C;Superfamily: cadherin; cadherin repeat homology
F;380-488/Domain: cadherin repeat homology <CR3>

Query Match 54.5%; Score 2379; DB 2; Length 884;
Best Local Similarity 53.5%; Pred. No. 1.3e-143;
Matches 471; Conservative 122; Mismatches 222; Indels 66; Gaps 13;

QY 9 ASLLLLQV-CWLOCAASEP--CRAVFREAEVTLAAGAEQEPQALGKV-FMGCPOGEPA 64
DB 10 ALLLLLQVSSWL-CQELPESCPGFSSEVTFPPVPERHLERGHVLRVFEFGCTGRPT 68
QY 65 LFSTDNDDFVRNGEIVQERSLK----- 88
DB 69 AFFSEDSRFKATDGTITVTKRHLKHLKLETSFLVRARDSSHRELSTKVTLSKMGHHHRH 128
QY 89 -----ERNP-LKIFPPSKRI-LRRHKRDWVAPISVPENGKGFPPQRLNQLKSNKDRDT 139
DB 129 HHRDPASESNPELLMFVSFVPLGRQRQKDWIPIPCPENEKEGFPKQLVQIKSNDRKET 198
QY 140 KIFYSITGQADSPGCVAFVEKETGMLLNKPLDREIEIAKELFGHAVSENGASVEDPM 199
DB 189 KVFYSITGQADKPPVGVFTIERETGWLKVTQPLDREAIAKYILSHAVSNGEAVEDPM 248
QY 200 NISIVTDNDHKPKFTQDFRGSVLEGLVPGTSMQVATDDEDDALTYNGVAVYSIHS 259
DB 249 EIVITVTDQNDNRPEFTQEFYEGSVAEGAVPGTSVMKVSATDADDVNTYNAIAIYIVS 308
QY 260 QEPKDPHDLMTIHRSTGTTISVTSGLDREKVPETLITIQATDMDGSGSTTAVAVVEIL 319
DB 309 QDELPKHKMFTVNRDGTGLSVITSLGLDRESYFTYTLVQVADLQEGSLSTTAKAVITVK 368
QY 320 DANDNAPMFDPOKYEAHVPENAVGHEVQRLTVTDLDPNSPAWRATYILMGDDGDGHFTI 379

Qy 140 KIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDREIEAKYELFGHAVSENGASVEDPM 199
Db 189 KVFYSITGQADKPPGVFIERETGMLKVQPLDREALAKYILYSHAVSSNGEAVEDEPM 248
Qy 200 NISIIVTDQNDHKPKFTQDTPRGSVLPGTSMQVMTATDEDAIYTVNGVVAIHS 259
Db 249 EIVITVTQDNDNRBEFTQVPVEGVAEGVSGTSMKVSAFADDDVNTYNAALAYTIVS 308
Qy 260 QEPKPDHLMETIHRSTGTSIVSSGLDREKVPYTLTIQATMDGSGSTTAVAVVEIL 319
Db 309 QDPFLPHKNMFTVARDTGIVSLVSGLDRESYPYTLVVQAADLQGGSLSTAKAVITVK 368
Qy 320 DANDNAPFDPQKYEAHPENAVGHEVQRLVTVDLAPNSPAWRATYLLIMGGDDGCHFTI 379
Db 369 DINDNAPVFNSTYQGVPENEVARATLKVTDADAPNTPAWKAIVTVV-NDPDQGFVV 427
Qy 380 TTHPESNOGILITTKGLDFFAKNQHTLYVEVTNAPVFLKLPSTATIVHVEDVNBAPV 439
Db 428 VTDPTTNDGILKTAKGLDFAKQYILHVRVENEPEPFGSLVPSTATVTVDDVDVNEAPI 487
Qy 440 FVPESKVEVOEGIPTCEPVCVYTAEPDK-ENOKISYRILRDPAGLAMPDPSGOVTAV 498
Db 488 FMPAERRVEVPEDFGVGOEITSYTAREPDTTMDQKITRIWRDTANWLEINPETGAIFTR 547
Qy 499 GTLDREDEQFVRNNIYEVWVLAMDNGSPPTTGTGTLTLIDVNDHGFVPPEPRQITICNQ 558
Db 548 AEMDREDAEHVKNSTYVALIATDGGSPATGTGTLVLLVLDVNDNAPIPEPRNMQFCQR 607
Qy 559 SPVRHLNITDKDLSPTSPFOALTDSDIYWTAEVNEEG-DTVVLSLKFLKQDYDV 617
Db 608 NPQPHIITLDPDLPNTPSPTELTGASVNWITIEYNDAAQESLILQPKRDLGEIKYKI 667
Qy 618 HLSLSDHCNKQLTIRATVCDCHVETCPGPKKGF-----ILPVLGAVTALLF 668
Db 668 HLLKADNQNQDQVTLTVHVCDCGTNNC---MKAGIVAAAGLVQPAILGLGILALLI 724
Qy 669 LLLVLLILVRKKRIKEPLLPEDDTRDNVYFVEEGGEDDQDYDTQLHRGLEARPEV 728
Db 725 LILLLLFLRRITVVKELLPDDTRDNVYVYDEEGGEDDQDFDLSQLHRGLDARPEV 784
Qy 729 VLRNDVAPTIPTMYRPRRANPEIGNFIENIKAANTDPTAPPYDTLLVDFYEGSGSD 788
Db 785 T-RNDVAPTLMSVQYRPRRANPEIGNFIENIKAANTDPTAPPYDLSLLVDFYEGSGSE 843
Qy 789 AASLSLTSASDODQDYDLNENSGRKKLADMYGGGEDD 829
Db 844 AASLSLNSBSDDQDYDLNENGNRPKLLADMYGGGEDD 884

RESULT 7
LJXLCP
EP-cadherin precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: B43785; A60128
R:Ginsberg, D.; Desimone, D.; Geiger, B.
Development 111, 315-325, 1991
A:Title: Expression of a novel cadherin (EP-cadherin) in unfertilized eggs and early xenopus embryos
A:Reference number: A43785; MUID:91372132; PMID:1893866
A:Accession: B43785
A:Molecule type: mRNA
A:Residues: 1-895 <GIN>
A:Cross-references: UNIPROT:P33148; UNIPARC:UPI0000171531; GB:X63720; NID:g64681; PIDN:C
A:Note: it is uncertain whether Met-1 or Met-16 is the initiator
R:Angres, B.; Mueller, A.H.J.; Kellermann, J.; Hausen, P.
Development 111, 829-844, 1991
A:Title: Differential expression of two cadherins in Xenopus laevis.
A:Reference number: A60128; MUID:91347911; PMID:1879345
A:Accession: A60128
A:Molecule type: protein
A:Residues: 171-177, 'I', 179-183, 'K', 185-189, 'XI' <ANG>
A:Cross-references: UNIPARC:UPI000017435C
A:Note: the material sequenced may have contained U-cadherin as well as E-cadherin

C;Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to be involved in many other cellular processes.
C;Keywords: cadherin; cadherin repeat homology
F;1-43/Domain: signal sequence #status predicted <SIG>
F;44-170/Domain: propeptide #status predicted <PRO>
F;171-895/Product: EP-cadherin #status predicted <MAT>
F;171-718/Domain: extracellular #status predicted <EXT>
F;173-278/Domain: cadherin repeat homology <CR1>
F;248-253/Region: cadherin binding #status predicted
F;281-391/Domain: cadherin repeat homology <CR2>
F;394-502/Domain: cadherin repeat homology <CR3>
F;503-610/Domain: cadherin repeat homology <CR4>
F;611-714/Domain: cadherin repeat homology <CR5>
F;719-743/Domain: transmembrane #status predicted <TM>
F;744-895/Domain: intracellular #status predicted <INT>
F;852-865/Region: serine-rich
F;440,696/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 54.4%; Score 2375; DB 1; Length 895;
Best Local Similarity 54.2%; Pred. No. 2.4e-143;
Matches 471; Conservative 128; Mismatches 214; Indels 56; Gaps 9;

Qy 11 LLLQVWCWQAASEPCRAVFEAEVTLFAGGAEPGQALGV-FMGCQGPALPSTD 69
Db 32 LCLLVVPSINADVSGCKPGFSAEYIFSVNRLELGRKLGKVNFSDCITRKHGLVDG 91
Qy 70 NDDFTV-----RNGETVQERRS 86
Db 92 DSFRVLDPGTVLVKRHVKLHKDTFTISTWDARGIKHSTNIAVAKRHSRGEAHSRSS 151
Qy 87 LKERNLKIPPSKRI-LRRHKRDWVAPISVPENKRGPFQRLNQLKSNKDRDTKIFYSI 145
Db 152 ---KLPLVLTPEHTGLKRRKRDWIPPIKVSENERGPFKRLVQIKSNKDRNKVYSI 208
Qy 146 TGGASDPPEGVFAVKECTGWLNLKPLDREETAKYELFGHAVSENGASVEDPMNLSIIV 205
Db 209 TGGADNPQGVFRIEWETGWLVTPLDREEDYKYVLSHAYSENGSPVEEPMETINV 268
Qy 206 TDQNDHKPKFTQDTPRGSVLPGTSMQVMTATDEDAIYTVNGVVAIHSQBPKDP 265
Db 269 IDQNDNRPKFTQDTPRGSVREGVQPGTQVMAVATDEDDNIDSLNGVLSILKQDPPEP 328
Qy 266 HDLMFTIHRSTGTSIVSSGLDREKVPYTLTIQATMDGSGSTTAVAVVEILDANDNA 325
Db 329 IPNLFTINRTGTVSLIGTGLDREKPEYTLTVQATDLEGAGLSVEGKAIITDANDNA 388
Qy 326 PMFDPPQYEAHPENAVGHEVQRLVTVDLAPNSPAWRATYLLIMGGDDGCHFTIHPES 385
Db 389 PIFDPKTYTALVPENEIGFEVQRLSVTDLMPGTPAWQAVYKIR-VNEGGFFNITDPES 447
Qy 386 NOGILTRKGLDPEAKNQHTLYVEVTNEAPFVLKLPSTATIVHVEDVNEAPVFPSPK 445
Db 448 NOGILTRKGLDPEAKNQHTLYVEVTNEAPFVLKLPSTATIVHVEDVNEAPFVPAVS 507
Qy 446 VVEYQEGIPTGPEVYVTAEDDPKEN-QKISYRILRDPAGLAMPDPSGOVTAVGTLDRE 504
Db 508 RVDVSEDLRGEKISILVAQDPKQIQKLSYFIGNDPARWLTVKNKNGIVTGNGLDRE 567
Qy 505 DEQFVRNNIYEVWVLAMDNGSPPTTGTGTLTLIDVNDHGFVPPEPRQITICNQSPRVHV 564
Db 568 SE-YVKNTYTVIMLVTDGVSVGTGTGLIHLVLDVNDNGVPSPRVFTMCDQNPSPQV 626
Qy 565 LNTDKDLSPHTSPFOALTDSDIYWTAEVNEEGDTVLSLKFLKQDYDVHLSLSDH 624
Db 627 LTISDADIPNTTYPYKVSLSHGSDLTWKAEDESKGTSMLLSPTQQLKGGYSIVVLSDA 686
Qy 625 GNKEQLTVIRATVCDCHGVETCPGPKKGFILP----VLGAVLALLFLLLVLLVLRKX 680
Db 687 QNNPQLTVVNAVTVCSCEGKAIKQEKLVGGFDLPILVILGSLVALLILFLLLVLLFKRK 746
Qy 681 RKIKEPILLPDEDDTRDNVYFVEEGGEDDQDYDTQLHRGLEARPEVLRNDVAPTIIP 740
Db 747 KVKKEPILLPDEDDTRDNVYFVEEGGEDDQDYDLSQLHRGLSRPD-IMRNDVVPITMP 805

Qy	741	TPMYRPPNPANPDEIGNFIENLKAANTDTPAPPYDTLLVDFYEGSGSDAASLSLSSAS	800
Db	806	APHYRPPSPNPDEIGNFIENLKAANTDTPAPPYDTLLVDFYEGSGSEAASLSLSSNS	865
Qy	801	DQDDQDYLVNKGWSRFFKLAADMYGGGGDD	829
Db	866	NDEHDYLVNKGWSRFFKLAADMYGGGGDD	894

RESULT 8
S43064

cadherin - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Aug-1999
C/Accession: S43064; S43065
R;Kuehl, M.
submitted to the EMBL Data Library, March 1994
A;Reference number: S43064
A;Accession: S43064
A;Molecule type: mRNA
A;Residues: 1-905 <KUE>
A;Cross-references: UNIPARC:UPI00001715C5; EMBL:X78546; NID:g468816; PIDN:CAA55292.1; PI
R;Herzberg, F.; Wildermuth, V.; Wedlich, D.
Mech. Dev. 35, 33-42, 1991
A;Title: Expression of Xbcad, a novel cadherin, during oogenesis and early development c
A;Reference number: S43065; MUID:92062581; PMID:1840622
A;Accession: S43065
A;Molecule type: mRNA
A;Residues: 'NSA', 462-697, 'O', 699-807, 'A', 809-840, 'N', 842-877, 'N', 879-883, 'N', 885-902, 'E'
A;Cross-references: UNIPARC:UPI0000171511; EMBL:X78546
C;Superfamily: cadherin; cadherin repeat homology
C;Keywords: calcium binding; cell adhesion; glycoprotein; transmembrane protein
F;181-286/Domain: cadherin repeat homology <CR1>
F;289-399/Domain: cadherin repeat homology <CR2>
F;402-510/Domain: cadherin repeat homology <CR3>
F;511-618/Domain: cadherin repeat homology <CR4>
F;619-722/Domain: cadherin repeat homology <CR5>

Query Match 54.3%; Score 2371; DB 2; Length 905;
Best Local Similarity 54.1%; Pred. No. 4, 4e-143;
Matches 469; Conservative 125; Mismatches 222; Indels 51; Gaps 8;

Qy	11	LLLLQVCWLQCAASEPCRAVFREAETLEAGGAEPQEQALGKY-FMCGCPGQBPALFSTD	69
Db	39	LCLLVVPSINVDVSGCQPGCFSSANYTFSVNRRELERGKLGKLVDCITRKHGLYDVG	98
Qy	70	NDDFTVNGETV-----QERRSLK	88
Db	99	DSRFRLVPLDGTVLVKRHVGLKSHKSDTRFTISTWDARGIKHSTNISVNVKRRHSGEEARS	158
Qy	89	ERNPDKTFPPKRI-LRRHKRDWVAPISVPENGGKPPQRLNOLKSNKDRDTKIFYSITG	147
Db	159	SELPVLTFPPSEKHTGKPKKRDWVPIPKVSENERGPPPKRLVQIKSNKEKLSKYSFYSITG	218
Qy	148	PGADSPPEGVFAVEKETGWLNLKPLDREIAKYELFPGHVAVSENGASVEDPMNISIIIVTD	207
Db	219	QGADTPEGIFRIEKETGMQVTRPLDREVEKYVLLSHAVSENGASVEEPMETITVID	278
Qy	208	QNDHKPKFTQDTRFGSLVGLVPGTSMQVATDEDDAIYTYNGVAVSYTHSQEPKDPHD	267
Db	279	QDNRRPKFTQPVFRGSRVREGVQPGTKVMSVSATDDDDSIDSLNGVIAYSILKQDPEPIP	338
Qy	268	LMFTIHRSTGTIVSISGLDREKVPYTLTIQATDMDGDGSTTTAVAVVEILDANDNAPM	327
Db	339	NLFTINRETGVSILIGLDREKFPETLIVQAADLOGAGLTAEGRKAVIEITDANDNAPI	398
Qy	328	FDPOKYEAHVPENAVGHVEQRLTVTDLDAPNSPAWRATYILMGDDGDGHFTITTHPESNQ	387
Db	399	FDPKTYTALVPENVEGVQRLSVTDLDMPGTAQWQVYKIR-VNEGGFNITTDPESNQ	457
Qy	388	GILTRTKGLDFEAKNQHTLVVEVTNEAPFVLKLPSTSTATIVVHVEDVNEAPVFPVPSKV	447


```
Qy 519 LAMNGSPPTTGTCTLLTLIDVNDHGVPEPRQITTCNOSPVRHVLNITDKDLSPTSP 578
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 LATDGGSPPTTGTCTLLTMDINDHGVPEPRQITTCNOSPVPQVNLITDKDLSPTAP 240

Qy 579 FQAQLTDDSDIYTAEVNEEGDVTVLSLKKFLKQDTYDVHLSLSDHGKKEQLTVIRATVC 638
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 FQAQLTHDSDVYTAEVNEKGDAVALSLKKFLKQGEYDVHLSLSDHGKKEQLTVIRATVC 300

Qy 639 DCHGHVETCPGWKGKGFILPVLGAVLALLFLLVLLVLLVLRKKIKKEPLLIPEDDTRDNV 698
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 DCHGNMVTCDPWTWGFLLPTLGAALALLLALLVLLVLRKKIKKEPLLIPEDDTRDNV 360

Qy 699 FYYEGGGEDQDYDITQLHRLGLEARPEVVLNDVAPTITPTMYRPRANPDEIGNFI 758
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 FYYEGGGEDQDYDITQLHRLGLEARPEVVLNDVAPSFIPTMYRPRANPDEIGNFI 420

Qy 759 IENLKAANTDPTAPPYDTLVDFYEGSGSDAASLSSTSSASQDQDYDLNMGWSRFXK 818
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 IENLKAANTDPTAPPYDLSLLVDFYEGSGSDAASLSSTSSDQDQDYNYNMGWSRFXK 480

Qy 819 LADMYGGGEDD 829
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 481 LADMYGGGQDD 491

RESULT 10
LUCGL
E-cadherin precursor, hepatic - chicken
N:Alternate names: L-CAM; liver cell adhesion molecule
C:Species: Gallus gallus (chicken)
C:Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: A30201; A29866; E29866
R:Sorkin, B.C.; Hemperly, J.J.; Edelman, G.M.; Cunningham, B.A.
Proc. Natl. Acad. Sci. U.S.A. 85, 7617-7621, 1988
A:Title: Structure of the gene for the liver cell adhesion molecule, L-CAM.
A:Reference number: A30201; MUID:89017248; PMID:3174655
A:Accession: A30201
A:Molecule type: mRNA
A:Residues: 1-81 <SOR>
A:Cross-references: UNIPROT:P08641; UNIPARC:UPI0000174355; EMBL:J04074
R:Gallin, W.J.; Sorkin, B.C.; Edelman, G.M.; Cunningham, B.A.
Proc. Natl. Acad. Sci. U.S.A. 84, 2808-2812, 1987
A:Title: Sequence analysis of a cDNA clone encoding the liver cell adhesion molecule, L-
A:Reference number: A29866; MUID:87204217; PMID:3472238
A:Accession: A29866
A:Molecule type: mRNA
A:Residues: 51-887 <GAL>
A:Cross-references: UNIPARC:UPI0000174356; EMBL:M16260
A:Accession: B29866
A:Molecule type: protein
A:Residues: 161-172;323-336;386-407;533-551 <GA2>
A:Cross-references: UNIPARC:UPI0000174357; UNIPARC:UPI0000174359;
C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
C:Genetics:
A:Introns: 23/3; 62/1; 138/1 183/3; 235/3; 284/1; 342/3; 385/3; 447/3; 529/2; 577/1; 652
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; liver; transmembr
F:1-26/Domain: signal sequence #status predicted <SIG>
F:161-887/Product: E-cadherin, hepatic #status experimental <PRO>
F:161-704/Domain: extracellular #status predicted <EXT>
F:163-268/Domain: cadherin repeat homology <CR1>
F:238-243/Region: cadherin binding #status predicted
F:271-381/Domain: cadherin repeat homology <CR2>
F:384-493/Domain: cadherin repeat homology <CR3>
F:494-601/Domain: cadherin repeat homology <CR4>
F:602-704/Domain: cadherin repeat homology <CR5>
F:705-735/Domain: transmembrane #status predicted <TM>
F:736-887/Region: intracellular #status predicted <INT>
F:291,346,564,643/Binding site: carbohydrate (Asn) (covalent) #status experimental
```

```
Query Match 52.0%; Score 2274; DB 1; Length 887;
Best Local Similarity 52.7%; Pred. No. 6.6e-137; Indels 55; Gaps 13;
Matches 461; Conservative 122; Mismatches 236;

Qy 8 LASLLLVQVCWLOCAASEPCRAVFRAEVTLLEAGAEQEPGQALGKV-FWGCGQGPALF 66
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 16 LVLLLLVQVGRRCDEAACPQGFAAETFSFSPQDPSVAAGRELGRVSAACSGRPWAVY 75

Qy 67 -STD-----NDQ-----FTVRNGETVQERRS----- 86
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 76 VPTDTRFKVNGDGVSTKRPPLTYLGRKKISFTIYAQAMGKRHSARVTVGRHRRHHNH 135

Qy 87 -LKERNPLKI-FP--SKRILRRHKRDWVAPISVPENKGGPPQRLNQLKSNKDRDTKIF 142
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136 HLQDTPPAVLTPFKHPDGLRQKRDWVIPPISCLNHRGPYPMRLVQIKSNKDKSKVY 195

Qy 143 YSTGTGASDPPEGVFAVEKETGWLKLLNKPDLREEIAKYELFGHVAENSEGASVEDPMNIS 202
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 196 YSITGGQADSPPVGIFIERETGWLVEVTEQLDREKIDRYTLLSHAVSASGQPVDEPMEI 255

Qy 203 IIVTDQNDHKPKFTQDTERGSVLEGLPGTSVMQVATDEDDAIYTVGVVAYSIIHQEP 262
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 256 ITVMDQNDKPKVFIEKVFVGYIEENAKPGTSVMVNTADADDVNTDNGIVSIVSQQP 315

Qy 263 KDPHDLMTIHRSTGTISVISGLDREKVPEYTLTIQATMDMGDGGSTTTAVAVEILDAN 322
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 316 PREHPQWFTIDPAKGIISVLGTGLDRETPNYTLIVQATDQEGKGLSNTAIIIEVTAN 375

Qy 323 DNAMPDPQKYEAHVPENAVGHEVQRLTVTDLDAPNSPAWRATYILMGDDGDGHFTITTH 382
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 376 DNIPIFNPTMYEGWEEENKPGTEVARLTVTDQDAPGSPAWQAVYHIKSGNLGDAFSITD 435

Qy 383 PESNOGILTRKGLDFEAKNQHTLYVEVTNEAPFVLKLPSTSTATIVVHVEDVNEAPVFP 442
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 436 PSTNNGILTKAGLDYETSKYDLVVTVENKVPSPITLSTASVLTVLDVNEPPVFP 495

Qy 443 PSKWVEVQEGIPGPEVCVVTAEADPKKE-NQKISYRILRDPAGLWAMDPSGOVTVAGTL 501
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 496 PIKRVGVPEDLPVGOQVTSYTAQDPDRDMRQKITRMSGSDPAGWLYIHPENGIVTATQPL 555

Qy 502 DREDEQFVRNNIYEVWVLMNDNGSPPTGTGTLTLLTLDVNDHGVPEPRQITICNOSP 561
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 556 DRESVHAI-NSTYKAILAVDNGIPDITGTGTLTLLQLQDNDNGPTPEPRSFICSRQPE 614

Qy 562 RHYLNIITDKDLSPTSPQALTDSDIYVTAENVEGDTVLSLKKFLKQDTYDVHLSL 621
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 615 KQILSIVDKDLPPTPTPFKAALHSGSNWTVIRGQ-DELA MGLKKELEPGEYINPVKL 673

Qy 622 SDHGNKEQLTVIRATVCDCHGHVETCP--GPWKGGF-----ILPVLGAVLALLFLVLL 675
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 674 TDSQGAQVTVQKVAQVCECEGTAKNCERRSYIVGGLGVPAIILGILGAILALLL 733

Qy 676 LVKKKKIKKPELLLPEDDTRDNVYVYEGEGGHEDDQYDITQLHRLGLEARPEVVLNDVA 735
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 734 FARRRKEVEPLLPEDDMDRDNVYVYDEEGGSEDDYDLSQLHRLGLDARPEVI-RNDVA 792

Qy 736 PTIIPTPMYRPRANPDEIGNFIENLKAANTDPTAPPYDTLVDEYEGSGSDAASLSL 795
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 793 PPLMAAPQYRPRANPDEIGNFIDENLKAANTDPTAPPYDLSLLVDFYEGGSEATSLSL 852

Qy 796 TSSASQDQDYDLNMGWSRFXKLADMYGGGEDD 829
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 853 NSSASQDQDYDLNMGWNRFXKLAEIYGGGEDD 886

RESULT 11
S47518
cadherin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 01-Feb-1995 #sequence revision 10-Nov-1995 #text_change 17-Mar-1999
C:Accession: S47518
R:Tooi, O.; Fujii, G.; Tashiro, K.; Shikawa, K.
Biochim. Biophys. Acta 1219, 121-128, 1994
```


A:Title: Molecular cloning of cDNA for XTCAD-1, a novel Xenopus cadherin, and its expression in the developing nervous system
A:Reference number: S47518; MUID:94368839; PMID:8086449
A:Accession: S47518
A:Molecule type: mRNA
A:Residues: 1-871 <TOO>
A:Cross-references: UNIPARC:UPI0000177AF8
C:Superfamily: cadherin; cadherin repeat homology
F:151-256/Domain: cadherin repeat homology <CR1>

Query Match	50.2%	Score 2192.5	DB 2	Length 871	
Best local Similarity	51.9%	Pred. No.1e-131			
Matches	460	Conservative 126	Mismatches 227	Indels 73	Gaps 16
Qy	1	MGLPR----	GPLASILLLOVC----	WLOCAASEPCRAVFEAEVTLEAGGAEQGOAL	51
Db	1	MGLKRPWLLGAVLLTLIOVGGLAETQ----	CRNGFSKEKYSEFLV-PKNLETKAL	53	
Qy	52	GKV-FMGCPCQEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIP-----			96
Db	54	GRVIFNCEGPVRIOFASKDPNPEIHKDGTVYIKNPAKMKDNKRTFRVLAWETKGHVST			113
Qy	97	-----	PSKRILRRHKRDWVAPISVPENKGGKPPFORLNOLK		132
Db	114	NITLKREGRHRODLFSKGKSHHPKSTGLKQKRDWVPIIYSENEKGFPPKRIYQIK			173
Qy	133	SNKDRDTKIYFSITPGCADSPPEGVAFEKGTGMLLNKPLDREIIAKYELFGHAVSENG			193
Db	174	SSYAKEVKVYISITGGQADTPPEGVFAIGREDGWNVTRPLDREDAINYVLFSHAVSSNG			233
Qy	193	ASVEDPMNISIIVTDQNDHKPKFTQDTFRGSVLGVLPGTISVMQVTTATDEDDAIYTVNGV			252
Db	234	ANVEDPMELIIVKQDQNDNPVFTQSVFEGSVPSGKEGTAVMTVSATDADDSVDVMTNGV			293
Qy	253	VAYSIHQEPKDPHDLMTFTHIRSTGTTISVASSGLDREKVPBYTTIQTATMD-GDGGTTT			311
Db	294	ITYSILNQEPKEPTNKMTFTHISEGLISVLTLGLDREKNPVYTLTIQAADGEFGKDRTTT			353
Qy	312	AVAVEITLDANDNAMPEDPKQYEAHVAPENAVGHVEQRLTVTTDLDPNPSPAWRATYLIIMG			371
Db	354	ATALIVMDTNDNPVFPDPTQYAKVPENEVGVEVARLTVTDEIEGTDANNAVYKIIKG			413
Qy	372	DGDGHFTTTHTPESNOGILTRKGLDPEAKNQHTLYVEVTNEAPFVLKLPSTSTATIVVHV			431
Db	414	NEANYFSIQT-DTGNIGLLTKVGLDYELKKQYILSVITNKANFSVPLQTSSTATVTVSV			472
Qy	432	EDVNEAPVFPVPSKVVEQEGIPGEBPVCVVTAEDDPKE-NQKISYRIILDDPAGLWMDP			490
Db	473	EDVNEAPIFLPVPKEVSVESDLPSPGVVATVTAQDPDKEQNKITYYIGNDPSRWVSNK			532
Qy	491	DSGOVTAVGTLDREDEQFVRNIIYEVVWLAMDNGSPPTGTGTLTLLTLIDVNDHGGVPPEP			550
Db	533	DNGIIVTGNLNDRE-SKPVLNNTYKVIILAADSGSPSATGTGTLVLNLLDVNDNGPFLPEP			591
Qy	551	RQITICNQSPVRHVNLNTDKDLSPTSPFOAQLTDDSDIYWTAEVNEBEGDTVWLSL--KK			608
Db	592	QQESFCQKDPGFRVFTIIDRLSPNTYPKAELTGESNENWTAIVTDK----ILLELPKK			647
Qy	609	FLKQDITVDVHLSLSDHGNKEQLTVIRATVCDCHGHVETC--PGPWKGF----ILPVIGA			662
Db	648	ELEITGVDMVTLLDLSFGLSNVTYKLIHTIICQCDGDKMQCEKAAIAGGLGISAIVGLIG			707
Qy	663	VLALLFLLVLLVLLAVRKRKIKPELLAPEDDTRONVFYEGEGGEGEDDYDITOLHRL			722
Db	708	ILALLLLLLLLLLFVRKKVVKPELLPEDETRONVFSYDEEGEGGEDQFDLSQLHRL			767
Qy	723	EARPEVLRNDVAPTIPTPMYRRPRPANPDEIGNFIIENLKAANTDPTAPPYDTLLVPDY			782
Db	768	DAREPVI-RNDVAP-VLAAQYRRPRPANPDEIGNFIDENLNAANDPTAPPYDLSLVPDY			825
Qy	783	EGSGSDAASLSLTSASDQDDYDYLNEGWSRPFKLADMYGGED			828
Db	826	EGSGSEAASSLNNSNDLDODYSALNDWGPRTKLADMYGGED			871

RESULT 12
TJTHUCN

1JHUCN

cadherin 2 precursor - human

N:Alternate names: N-cadherin; neuronal cadherin

C:Species: Homo sapiens (man)

C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004

C:Accession: A38870; S11487; JQ0751; S13799

R:Reid, R.A.

submitted to the EMBL Data Library, November 1990

A:Reference number: A38870

A:Accession: A38870

A:Molecule type: mRNA

A:Residues: 1-906 <REI>

A:Cross-references: UNIPROT:P19022; UNIPARC:UPI000016ADB; EMBL:X54315; NID:g34998; PIDN:249998

R:Reid, R.A.; Hemperly, J.J.

Nucleic Acids Res. 18, 5896, 1990

A:Title: Human N-cadherin: nucleotide and deduced amino acid sequence.

A:Reference number: S11487; MUID:91016946; PMID:2216790

A:Accession: S11487

A:Molecule type: mRNA

A:Residues: 1-340, 'N', 342-698, 'R', 700-704, 'F', 706-906 <RE2>

A:Cross-references: UNIPARC:UPI000017434F; EMBL:X54315

A:Note: this sequence has been revised in reference A38870

R:Walsh, F.S.; Barton, C.H.; Putt, W.; Moore, S.E.; Kelsell, D.; Spurr, N.; Goodfellow, I.

J. Neurochem. 55, 805-812, 1990

A:Title: N-cadherin gene maps to human chromosome 18 and is not linked to the E-cadherin

A:Reference number: JQ0751; MUID:90347462; PMID:2384753

A:Accession: JQ0751

A:Molecule type: mRNA

A:Residues: 160-194, 'IR', 197-211, 'L', 213-227, 'Q', 229, 'N', 231-235, 'G', 237-248, 'T', 250-356, 'G'

A:Cross-references: UNIPARC:UPI000016ADC0; GB:M34064

C:Comment: Cadherins mediate calcium-dependent intercellular adhesion and are thought to

C:Genetics:

A:Gene: GDB:CDH2; NCAD

A:Cross-references: GDB:128185; OMIM:114020

A:Map position: 18q12.1-18q12.1

C:Superfamily: cadherin; cadherin repeat homology

C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane protein

F:1-27/Domain: signal sequence #status predicted <SIG>

F:28-159/Domain: propeptide #status predicted <PRO>

F:160-906/Product: N-cadherin #status predicted <MAT>

F:160-714/Domain: extracellular #status predicted <EXT>

F:162-267/Domain: cadherin repeat homology <CR1>

F:237-242/Region: cadherin binding #status predicted

F:270-382/Domain: cadherin repeat homology <CR2>

F:385-497/Domain: cadherin repeat homology <CR3>

F:500-605/Domain: cadherin repeat homology <CR4>

F:606-712/Domain: cadherin repeat homology <CR5>

F:715-746/Domain: transmembrane #status predicted <TM>

F:747-906/Domain: intracellular #status predicted <INT>

F:865-878/Region: serine-rich

F:130,273,325,402,572,622,651,692/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	39.1%	Score 1707;	DB 1;	Length 906;
Best Local Similarity	42.7%	Pred. No. 9.6e-101;		
Matches 390; Conservative 129; Mismatches 292; Indels 102; Gaps 27;				
Qy	6	GPLASILLLOVCMLOCAASEP----	CRAVFRAEAVTLEAGGAEOEQALGKV-FMGCP	59
Db	6	GAULTLLPLLALLQASVEASGEIALCKTGFP-	DVYSAVLSKDVHSGQPLLVKFN	64
Qy	60	GQ-----EPALFSTDND--FTVRN-----	GETVOER-----	84
Db	65	GKRKVQYSESEPADFKVDEGMYAVRSFPLS	SEHAKFLIYAQDKETQEKWQAVKLSLK	124
Qy	85	-----RSLKERNPLK--IFP---SKRI--	LRRHKRDWVAPISVPENKGKGPFPORLNQK	132
Db	125	PTTEESVKESAEEVEIIFPRQSKSGHQRQKRDW	IPINLINSRSGPFPQELVRI	184
Qy	133	SNKDRDTKIFYSITGPGADSPFPGVFAVEKETG	WLLNKLPLDREETIAKLYELFGHVS	192
Db	185	SDRKNLISLRYVTGPGADPPGTFINPISGOLSV	TKPLDREOIRARFLRAHAVDING	244

QY 814 SRFKLADMYGGGED 828
|||||
Db 898 PRFKLADMYGGGD 912
|||||
RESULT 14
IJM5CN
N-cadherin precursor, neuronal - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: A32759; A46163
R: Miyatani, S.; Shimamura, K.; Hatta, M.; Nagafuchi, A.; Nose, A.; Matsunaga, M.; Hatta,
Science 245, 631-635, 1989
A:Title: Neural cadherin: role in selective cell-cell adhesion.
A:Reference number: A32759; MUID:89346748; PMID:2762814
A:Accession: A32759
A:Molecule type: mRNA
A:Residues: 1-906 <MI>
A:CROSS-references: UNIPROT:P15116; UNIPARC:UPI000002924B; GB:M31131; NID:g192327; PIDN:
R: Miyatani, S.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Takeichi, M.
Proc. Natl. Acad. Sci. U.S.A. 89, 8443-8447, 1992
A:Title: Genomic structure and chromosomal mapping of the mouse N-cadherin gene.
A:Reference number: A46163; MUID:92409532; PMID:1528849
A:Accession: A46163
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 839-906 <MI2>
A:CROSS-references: UNIPARC:UPI0000170C39; GB:S45011; NID:g256010; PIDN:AAB23356.1; PID:
A:Note: sequence extracted from NCBI backbone (NCBIN:113759, NCBI:P.113760)
C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
C:Superfamily: cadherin; cadherin repeat homology
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-159/Domain: propeptide #status predicted <PRO>
F:160-906/Product: N-cadherin #status predicted <WAT>
F:160-714/Domain: extracellular #status predicted <EXT>
F:162-267/Domain: cadherin repeat homology <CR1>
F:237-242/Region: cadherin binding #status predicted
F:270-382/Domain: cadherin repeat homology <CR2>
F:385-497/Domain: cadherin repeat homology <CR3>
F:500-605/Domain: cadherin repeat homology <CR4>
F:606-714/Domain: cadherin repeat homology <CR5>
F:715-746/Domain: transmembrane #status predicted <TMW>
F:747-906/Domain: intracellular #status predicted <INT>
F:865-978/Region: serine-rich
F:190,273,325,402,572,651,692/Binding site: carbohydrate (Asn) (covalent) #status predic

Db 306 YRILSQAPSTPSPNNMFTINNETGDIITVAAGLDREKVOQYTLIIQATMEGNTYGLSNT 365
QY 312 AVAVVEILDANDNAPMFDPOKYEAHVPENAVGHEVORLTVTDLDPNSPAWRATYILMGG 371
|||||
Db 366 ATAVITVDNDNPPPEFTAMTFYGEVPENRVDIVANLTVTDKQPHTPANNAAYRISGG 425
QY 372 DGDHFTITTHPESNOGILTRKGLDFEAKOHTLVVEVTNEAPFVLKL---PTSTATIV 428
|||||
Db 426 DPTGRFAILTDSNNDGLVTWVKPIDFETNRMFVLTVAAENOVPLAKGIHQHPQSTATVS 485
QY 429 VHVEDVNEAPVFPVPSKVVEGEGIPTGEPVCVYTAEDPK-ENOKISVRIILDRPAGWLA 487
|||||
Db 486 VTVIDVNEPYPAPNPKIIRQEGELHAGTMLTTLTAQDDPDRIYQQNIRYTKLSDPANWLK 545
QY 488 MDPDSQVTVAGTLDREDFQVRNNIYEVNVLAMDNQSPPTGTGTLLTLTLDVNDHGPV 547
|||||
Db 546 IDPVNGQITTIADLRE-SPYVQNNIYNATFLASDNGIIPMSGTGLQIYLLDINDNAPQ 604
QY 548 PEPRQITICNOQSPVRHVLNIT--DKDLSPTSPFQAQLTDDSDIY---WTAEVNE-EGDT 601
|||||
Db 605 VLPQEAETC-ETPEPNSINIAALVDIDNAGPFADPLPLSPVTIKRNTW--INRLNGDF 661
QY 602 VVLSLK-KELKQDYDVHLSLSDHGN--KEQLTVIRATVCDCHG-----VETCPGPWK 653
|||||
Db 662 AOLNLKIKLEAGIYEVPIIITDSGNPPKSNISILRVKVCQDSNGDCTDVIDRIVGAGLG 721
QY 654 GFILPVLGAVLALLFLLLVLLLV-----RKRKIKEPLLLPEDDTRONVFFYGE 703
|||||
Db 722 -----TGAIALLLCIIILLVLMFVVMKROKROKQALLIDPEDDVRDNLKYDE 775
QY 704 EGGBEDQYDITOLHRLGLEARPEV-----VLKNDVAPITPTPMYRPPA--NPDEIGN 756
|||||
Db 776 EGGBEDQYDLSQLQPDVTBPAIKVGIIRRLDERP-IHAEPQYFVRSAAHPHGDIGD 834
QY 757 FIENLKAANTOPTAPPYDTLLVFDYEGSGSDAASLSLTSASDQDQDYVYLNWGSRF 816
|||||
Db 835 FINEGLKAADNPTAPPYDLSLLVFDYEGSGTAGSLSSSSSGGQDQDYDLDNDGPRF 894
QY 817 KKLADMYGGGED 828
Db 895 KKLADMYGGGD 906
|||||
RESULT 15
IJCHCR
R-cadherin precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: JH0424
R: Inuzuka, H.; Miyatani, S.; Takeichi, M.
Neuron 7, 69-79, 1991
A:Title: R-cadherin: a novel Ca2+-dependent cell-cell adhesion molecule expressed in the
A:Reference number: JH0424; MUID:91299341; PMID:1712604
A:Accession: JH0424
A:Molecule type: mRNA
A:Residues: 1-913 <INU>
A:CROSS-references: UNIPROT:P24503; UNIPARC:UPI0000126D94; GB:D14459; GB:D00849; NID:g42;
A:Experimental source: retina
C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
C:Superfamily: cadherin; cadherin repeat homology
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-166/Domain: propeptide #status predicted <PRO>
F:167-913/Product: R-cadherin #status predicted <WAT>
F:167-721/Domain: extracellular #status predicted <EXT>
F:169-274/Domain: cadherin repeat homology <CR1>
F:244-249/Region: cadherin binding #status predicted
F:277-389/Domain: cadherin repeat homology <CR2>
F:392-504/Domain: cadherin repeat homology <CR3>
F:507-612/Domain: cadherin repeat homology <CR4>
F:613-721/Domain: cadherin repeat homology <CR5>
F:722-753/Domain: transmembrane #status predicted <TMW>
F:754-913/Domain: intracellular #status predicted <INT>

Job time : 48 secs

F;870-885/Region: serine-rich
F;280,409,554,629,658,699/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match		38.8%;	Score 1694;	DB 1;	Length 913;
Best Local Similarity		41.5%;	Pred. No. 6.5e-100;		
Matches 383;		Conservative 135;	Mismatches 286;	Indels 118;	Gaps 25;
Qy	11	LLLLQVCWLQCAA-----SEPCRAVFRAEAVTELEAGGAEQEPGQALGKV-FMGCPGOE	62		
Db	7	LLLLVLLVWGSAALNGDLTVRPTCKPGFSEDEYTAFSQNIIME-GQKLLKVKFNKNCAGNK	65		
Qy	63	PALFSTDNDDDFTVR-----	76		
Db	66	GVRVETNSLDFKVRADGTWVAHVQVMASKQLILMVTAWDPQTLGRWEAIVRPLVGEKLQ	125		
Qy	77	-NGETVQERRS-----LKERNPLKIFP-----SKRILRRHKRDWVAIPISVPENCKGPFPO	126		
Db	126	HNGHKPKGRKSGPVDLAQQOQSDTLTPWRQHQSAGLRQRKRDWVIPPINVPENSRGPFPO	185		
Qy	127	RLNOLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDREEIAKYELFGH	186		
Db	186	QLVRIKDKKEIHRIISITGVGADQPPMEVFSIDPVGSMYVTRPMDREERASYHLRAH	245		
Qy	187	AVSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGSVLEGVLPGTSVMQVTTATDEDDAI	246		
Db	246	AVDMNGNKVENPIDLIYIIVIDMNDNRPEFINQVYNGSVDEGSKPGTYVMTVTANDADDST	305		
Qy	247	YTYNGVAVAYSHSQEPKDPHDLMETIHRSTGTSIVISSGLDREKVPETILTIOATDMGD	306		
Db	306	-TANGMYRYRIVTQTPQSPSQNMFTINSETGDIYVTAAGLDREKVQOQVMYIVQVATDMEGN	364		
Qy	307	--GSTTTAVAVRILDANDNAPMFDPKYEAHVPENAVGVHQRVLTVDLDAFNSPAWR	363		
Db	365	LNYSGLSWTATAIIVTDVNDNPPEFTTSTYSGEVPENRVEVVANLIVMDRDQPHSPWN	424		
Qy	364	ATYLMGGDDGDHFTITTHPESNOGIITFRKGLDFEAKNQHTLYVEVTNEAPFVLKLP--	421		
Db	425	AIYRIISGDPGSGHTTIRTPVTNEGTVVKAVDYENMRAFMLTVMVSNQAPLASGIQMS	484		
Qy	422	-TSTATIVVHVEDVNEAPVFVPPSKVEVEGIGITGEPVCVYTAEDPDK-ENKISYRIL	479		
Db	485	FQSTAGVTISVTDVNEAPYFPTNHKLIRLEEGVPTGTLTTFSAVDPDFRPMQAVRYSKL	544		
Qy	480	RDPAWGLAMDPSGQVAVGLTDRDEQFVRNNIYEVWVLAMDNGSPPTGTGTGLLLTLI	539		
Db	545	SDPANWLINATNGQITTAAVLDRES-DYKNNVYEATFLAADNGIPPASGTGTLQIYLI	603		
Qy	540	DVNDHGPVPEPQITICNQSPVRHVNIT--DKDLSPTSPFQAQLTDDSDIY---WTAE	594		
Db	604	DINDNAPELLPKEAQC-EKENLVNINITAADADIDPNVGPFPPELPSPSAVRKNWTI-	661		
Qy	595	VNEEGDVTVLSLK-KFLKQDITYVHLSLDHGKNE--QLTVIRATVDCD--HGHVETCPG	649		
Db	662	TRLNGDYAQLSLRIMYLEAGYDYVPIIVTDSGNPPLYNTSIKVKVCPCDENGDCCTTGA	721		
Qy	650	PWKGGFILPVLGAVLALFLLLVLLLV-----RKKRKIKPEPLLPEDDDTRDNVF	699		
Db	722	VAAAGL---GTGAIAIILICIIILITWVLLFVVMWKREKERHTKQLIIDPEDDVRNII	778		
Qy	700	YYGEGGEEQDYDITQLHRGLEARPEV-----VLRNDVAPTIITPTMYRPRPA	749		
Db	779	KYDEEGGEEQDYDLSLQ-----QPETMDHVLNKAPGVRVRVERP-IGAEPQYPIRPV	832		
Qy	750	--NPEDEIGNFIENLKAANTDPTAPPYDTLLVDFEGSGSDAASLSLTSSASDQDQDYD	807		
Db	833	IPHEGDIIGDFINEGLRAADNDPTAPPYDLSLVDFYEGSGTAGSVSLNSSSG-DQDYD	891		
Qy	808	YLNWGSRSFKKLADMYGGGEDD	829		
Db	892	YLNDWGPFRFKKLADMYGGGEED	913		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2005, 09:01:57 ; Search time 189 Seconds
(without alignments)
1927.222 Million cell updates/sec

Title: US-10-696-639-39

Perfect score: 4369

Sequence: 1 MGLPRGLASLLLLQVWLQ.....NEWGSRFKLADMYGGSD 829

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4369	100.0	829	5 ABG61897	Abg61897 Prostate
2	4369	100.0	829	5 AB881476	Ab881476 Human P-c
3	4369	100.0	829	5 ABJ05598	Abj05598 Breast ca
4	4369	100.0	829	5 AAU97492	Aau97492 Human P-c
5	4369	100.0	829	5 AAM50864	Aam50864 Cadherin
6	4369	100.0	829	6 ABP58357	Abp58357 Human P-c
7	4369	100.0	829	7 ADC15497	Adc15497 Human bas
8	4369	100.0	829	7 ADD14190	Add14190 Human src
9	4369	100.0	829	7 ADN39482	Adn39482 Cancer/an
10	4369	100.0	829	7 ADN39578	Adn39578 Cancer/an
11	4369	100.0	829	7 ADN39545	Adn39545 Cancer/an
12	4369	100.0	829	7 ADN39465	Adn39465 Cancer/an
13	4369	100.0	829	8 ADL06565	Adl06565 Human tum
14	4369	100.0	829	8 ADQ20015	Adq20015 Human sof
15	4369	100.0	829	8 ADP26905	Adp26905 Human P-c
16	4369	100.0	829	9 ADV73198	Adv73198 Human col
17	4369	100.0	829	9 AEB87752	Aeb87752 Human P-c
18	4361	99.8	829	5 ABP54683	Abp54683 Metastati
19	4361	99.8	829	6 ABR58670	Abr58670 Human can
20	4361	99.8	829	6 ABU56434	Abu56434 Lung can
21	4361	99.8	829	7 ADN39020	Adn39020 Cancer/an
22	4361	99.8	829	7 ADN39928	Adn39928 Cancer/an
23	4361	99.8	829	8 ADL70229	Adl70229 Colon can
24	4361	99.8	829	8 ADN59621	Adn59621 Colon neo

25	4361	99.8	829	8 ADN03948	Adn03948 Antipeori
26	4361	99.8	829	8 ADO28657	Ado28657 Human CAD
27	4361	99.8	829	9 ADX83229	Adx83229 Human TEG
28	4351	99.6	829	6 ABU56670	Abu56670 Lung can
29	3643.5	83.4	700	8 ADP29426	Adp29426 Human sec
30	3643.5	83.4	700	8 ADP29333	Adp29333 Human sec
31	2420	55.4	882	2 AAY09375	Aay09375 Wild-type
32	2420	55.4	882	3 AAB35730	Aab35730 Human E-c
33	2420	55.4	882	4 AAB73490	Aab73490 Human E-c
34	2420	55.4	882	5 ABG96293	Abg96293 Human ova
35	2420	55.4	882	5 AAU78051	Aau78051 Human E-c
36	2420	55.4	882	5 ADL15660	Adl15660 Human E-c
37	2420	55.4	882	8 ADN03700	Adn03700 Antipeori
38	2420	55.4	882	8 ADP26906	Adp26906 Human E-c
39	2420	55.4	882	8 ABM81765	Abm81765 Tumour-as
40	2420	55.4	882	8 ADR67260	Adr67260 Human bla
41	2420	55.4	882	9 AEB87740	Aeb87740 Human E-c
42	2420	55.4	896	8 ADR66291	Adr66291 Human pro
43	2420	55.4	896	8 ADR66852	Adr66852 Human pro
44	2420	55.4	923	8 ADR66668	Adr66668 Human pro
45	2420	55.4	923	8 ADR66326	Adr66326 Human pro

ALIGNMENTS

RESULT 1
ABG61897
ID ABG61897 standard; protein; 829 AA.
XX
AC ABG61897;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #98.
DE
KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS Mammalia.
XX
PN WO200230266-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US032045.
XX
PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733288.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX
PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
XX
DR WPI; 2002-471335/50.
DR N-PSDB; ABK92214.
XX
XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
PS Claim 27; Page 384-385; 436pp; English.
XX
XX
CC The present invention relates to methods of detecting a prostate cancer-associated transcript in a cell from a patient. The method comprises

102 a

CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
XX
SQ Sequence 829 AA;

Query Match 100.0%; Score 4369; DB 5; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGLASLLLLQVCWLCQAASEPCRAVFAEVTLKAGGAEQEQALGKVFMCPCG 60
Db 1 MGLPRGLASLLLLQVCWLCQAASEPCRAVFAEVTLKAGGAEQEQALGKVFMCPCG 60

Qy 61 QEPALFSTDDNDFVRNGETVOERRSLKERNPLKIFPSKILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDDNDFVRNGETVOERRSLKERNPLKIFPSKILRRHKRDWVAPISVPENG 120

Qy 121 KGPPFQRLNQLKSKNRDTKIFYGITGADSPPEGVFAVEKETGWLNNKPLDREBIAK 180
Db 121 KGPPFQRLNQLKSKNRDTKIFYGITGADSPPEGVFAVEKETGWLNNKPLDREBIAK 180

Qy 181 YELFGHAUSENGASVEDPMNISIIVTDQNDHKPFTQDTRGSLVLEGVLPQTSVMQVAT 240
Db 181 YELFGHAUSENGASVEDPMNISIIVTDQNDHKPFTQDTRGSLVLEGVLPQTSVMQVAT 240

Qy 241 DEDDAIYNGVWVAYSISQBPDPHLMFTIHRSTGTISVISGLDREKVPETLTIQA 300
Db 241 DEDDAIYNGVWVAYSISQBPDPHLMFTIHRSTGTISVISGLDREKVPETLTIQA 300

Qy 301 TDMGDGSGTTTAVAVVEILDANDNAMPDPQKYEAHPVNAVHGVQLTVDLDAENSP 360
Db 301 TDMGDGSGTTTAVAVVEILDANDNAMPDPQKYEAHPVNAVHGVQLTVDLDAENSP 360

Qy 361 AWRTYILMGDDGDGHFTITTHPSNQILTRKGLDPEAKNQHTLYVEVNEAPFVLKL 420
Db 361 AWRTYILMGDDGDGHFTITTHPSNQILTRKGLDPEAKNQHTLYVEVNEAPFVLKL 420

Qy 421 PTSTATIIVHVEDVNEAPVFPVPPSKVVEVQEGIPTGEPVCVYTAEDDPKENOKISYRILR 480
Db 421 PTSTATIIVHVEDVNEAPVFPVPPSKVVEVQEGIPTGEPVCVYTAEDDPKENOKISYRILR 480

Qy 481 DPAGWLAMPDPSGQVAVGTLTLDREDEQFVRNNIYEVWVLAMDNQSPPTGTGTLTLLID 540
Db 481 DPAGWLAMPDPSGQVAVGTLTLDREDEQFVRNNIYEVWVLAMDNQSPPTGTGTLTLLID 540

Qy 541 VNDHGPVPEPQITICNSPVRHVLNITDKLSPTSFPQALTDSDSIYTAEVNEEGD 600
Db 541 VNDHGPVPEPQITICNSPVRHVLNITDKLSPTSFPQALTDSDSIYTAEVNEEGD 600

Qy 601 TVVLSLKKFLKQDITVDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFLLPVL 660
Db 601 TVVLSLKKFLKQDITVDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFLLPVL 660

Qy 661 GAVLALLFLLVLLVLRKGRKIKEPPLLPPEDDTRDNVFFYGEESGGEDQDITQLHR 720
Db 661 GAVLALLFLLVLLVLRKGRKIKEPPLLPPEDDTRDNVFFYGEESGGEDQDITQLHR 720

Qy 721 GLEARPEVLRNDVAPTIIPTPMYRPRPANDEIGNFIENLKAANTDPTAPPDITLVF 780
Db 721 GLEARPEVLRNDVAPTIIPTPMYRPRPANDEIGNFIENLKAANTDPTAPPDITLVF 780

Qy 781 DYEGSGSDAASLSLTSASDQDDYDLNWSGRFCKLADMYGGEDD 829

Db 781 DYEGSGSDAASLSLTSASDQDDYDLNWSGRFCKLADMYGGEDD 829

RESULT 2
ABBS1476
ID ABB81476 standard; protein; 829 AA.
XX ABB81476;
AC ABB81476;
XX
DT 30-AUG-2002 (first entry)
XX
DE Human P-cadherin protein SEQ ID NO:50.
XX
XX Cadherin; alpha-catenin; cancer; beta-catenin binding domain; melanoma;
XX O-catenin; colon cancer.
XX Homo sapiens.
XX OS
XX US2002045591-A1.
XX PN
XX 18-APR-2002.
XX PD
XX 17-JUL-2001; 2001US-00905983.
XX PF
XX 26-MAY-1998; 98IL-00124650.
XX PR
XX 26-MAY-1999; 99US-00318633.
XX
XX (SEIG/) GEIGER B.
XX PA (BENZ/) BEN-ZE'EV A.
XX PA (SADO/) SADOT E.
XX
XX Geiger B, Ben-Ze'ev A, Sadot E;
XX WPI; 2002-499105/53.
XX DR N-PSDB; ABN89393.
XX
XX New construct encoding soluble cytoplasmic portion of cadherin including
XX beta catenin binding domain useful in treating cancer associated with
XX high beta-catenin activity e.g. colon cancer and melanoma.
XX
XX Example 3; Page 53-55; 102pp; English.
XX
XX The present invention describes a pharmaceutical composition for treating
XX cancer associated with abnormally high beta-catenin activity. The
XX pharmaceutical composition comprises a gene therapy vehicle harbouring a
XX polynucleotide that contains: (i) a nucleotide sequence encoding a
XX soluble cytoplasmic portion of a cadherin which lacks a transmembrane
XX portion and an extracellular portion of the cadherin, and includes a beta
XX -catenin binding domain; and (b) an upstream promoter for directing
XX expression of the soluble cytoplasmic portion of the cadherin in a
XX mammalian cell. Also described is a pharmaceutical composition for
XX treating cancer associated with abnormally high activity levels of beta-
XX catenin comprising a gene therapy vehicle harbouring a polynucleotide
XX that contains: (a) a nucleotide sequence encoding an o-catenin; and (b)
XX an upstream promoter for directing expression of the o-catenin in a
XX mammalian cell. The pharmaceutical compositions have cytostatic activity
XX and can be used in the suppression of beta-catenin-mediated
XX transactivation. They can be used for treating cancers associated with
XX abnormally high activity levels of beta-catenin such as colon cancers and
XX melanomas, by reducing these high activity levels of beta-catenin in
XX mammalian cells. The present sequence represents human p-cadherin which
XX is used in the exemplification of the present invention
XX
SQ Sequence 829 AA;

Query Match 100.0%; Score 4369; DB 5; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGLASLLLLQVCWLCQAASEPCRAVFAEVTLKAGGAEQEQALGKVFMCPCG 60
Db 1 MGLPRGLASLLLLQVCWLCQAASEPCRAVFAEVTLKAGGAEQEQALGKVFMCPCG 60

QY	61	QEPALFSTNDNDDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG	120
Db	61	QEPALFSTNDNDDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG	120
QY	121	KGPPFORLNQKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREEIAK	180
Db	121	KGPPFORLNQKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREEIAK	180
QY	181	YELFGHVAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLGVLPGTSMVQVAT	240
Db	181	YELFGHVAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLGVLPGTSMVQVAT	240
QY	241	DEDDAIYTYNGVVAYSIHQSEPKDPHDLMTFTHRSTGTISVSSGLDREKVPYETLTIOA	300
Db	241	DEDDAIYTYNGVVAYSIHQSEPKDPHDLMTFTHRSTGTISVSSGLDREKVPYETLTIOA	300
QY	301	TMMDGSGSTTTAVAVVEILDANDNAPMFPDQKYEAHVPENAVGVHVEVQRLTVDLDPNSP	360
Db	301	TMMDGSGSTTTAVAVVEILDANDNAPMFPDQKYEAHVPENAVGVHVEVQRLTVDLDPNSP	360
QY	361	AWRATYLMGGDDGDHFTTTHPESNOGILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420
Db	361	AWRATYLMGGDDGDHFTTTHPESNOGILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420
QY	421	PTSTATIVVHVEDVNEAPVFPSPKVVVEQGIPTGEPVCVYTAEDPDKENQKISYRIILR	480
Db	421	PTSTATIVVHVEDVNEAPVFPSPKVVVEQGIPTGEPVCVYTAEDPDKENQKISYRIILR	480
QY	481	DPAGWLAMPDSCQVTAAGTLDREDFQVRNNIYEVWVLAMONGSPPTTGTGLTLTLTD	540
Db	481	DPAGWLAMPDSCQVTAAGTLDREDFQVRNNIYEVWVLAMONGSPPTTGTGLTLTLTD	540
QY	541	VNDHGVPVPRQITICNQSPVRHVLNITDKDLSPHTSFPQAOLTDSDIYWTAEVNEED	600
Db	541	VNDHGVPVPRQITICNQSPVRHVLNITDKDLSPHTSFPQAOLTDSDIYWTAEVNEED	600
QY	601	TVVLSLKFKLQDQYDVHLSDHGNKEQLTVIRATVCDCHGVETCPGPKGFGFILPVL	660
Db	601	TVVLSLKFKLQDQYDVHLSDHGNKEQLTVIRATVCDCHGVETCPGPKGFGFILPVL	660
QY	661	GAVALLLFLLVLLVLRKRIKEPILLPEDDTRONVIFYEGEGGEDDQDYDTQLHR	720
Db	661	GAVALLLFLLVLLVLRKRIKEPILLPEDDTRONVIFYEGEGGEDDQDYDTQLHR	720
QY	721	GLEARPEVLRNDVAPTIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLVLF	780
Db	721	GLEARPEVLRNDVAPTIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLVLF	780
QY	781	DYEGSGSDAASLSLTSASDQDDQDYDLNWSGRFCKKLADMYGGGEDD	829
Db	781	DYEGSGSDAASLSLTSASDQDDQDYDLNWSGRFCKKLADMYGGGEDD	829

RESULT 3

ABJ05598
ID ABJ05598 standard; protein; 829 AA.

XX
AC ABJ05598;

XX
DT 14-NOV-2002 (first entry)

XX
DE Breast cancer-associated protein 63.

XX
KW Breast cancer; breast cancer-associated gene sequence; drug development;
KW pharmacogenetics; biosensor development.

XX
OS Unidentified.

XX
PN WO200259377-A2.

XX
PD 01-AUG-2002.

XX

PF	24-JAN-2002;	2002WO-US002242.	
XX	24-JAN-2001;	2001US-0263965P.	
PR	02-FEB-2001;	2001US-0265928P.	
PR	09-APR-2001;	2001US-00829472.	
PR	09-APR-2001;	2001US-0282698P.	
PR	04-MAY-2001;	2001US-0288590P.	
PR	29-MAY-2001;	2001US-0294443P.	
XX	(EOSB-) EOS BIOTECHNOLOGY INC.		
PA	Mack DH, Gish KC, Afar D;		
XX	WPI; 2002-583738/62.		
DR	N-PSDB; ABT07755.		
XX	Detecting a breast cancer-associated transcript in a patient's cell,		
PT	useful for diagnosing breast cancer, comprises contacting a biological		
PT	sample with a polynucleotide that selectively hybridizes with breast		
PT	cancer nucleic acids.		
XX	Disclosure; Page 405; 414pp; English.		
PS	The invention comprises a method of detecting a breast cancer-associated		
XX	transcript in a cell from a patient. The method of the invention involves		
CC	contacting a biological sample from the patient with a nucleotide that		
CC	hybridizes to one of the 69 breast cancer-associated gene sequences shown		
CC	in the specification. The method of the invention is useful in the		
CC	diagnosis or prognosis of breast cancer, and for detecting genes that are		
CC	up or down-regulated in breast cancer cells. Genes identified by the		
CC	method of the invention can be used in diagnostic purposes and also as		
CC	targets for screening for therapeutic compounds that modulate breast		
CC	cancer (e.g. hormones or antibodies). Identification of genes that are		
CC	over or under expressed in breast cancer can additionally provide high-		
CC	resolution, high-sensitivity datasets which can be used in the areas of		
CC	diagnostics, therapeutics, drug development, pharmacogenetics, protein		
CC	structure and biosensor development. Amino acid sequences ABJ05536 -		
CC	ABJ05604 represent the proteins encoded by the 69 breast cancer-		
CC	associated genes of the invention		
XX	Sequence 829 AA;		
SQ	Query Match	100.0%; Score 4369; DB 5; Length 829;	
	Best Local Similarity	100.0%; Pred. No. 0;	
	Matches 829; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	MGLPRGPLASLLLVCMVLCQAASPCRAVFEAEVTLKAGGAEQPGQALGVFMGCPG	60
Db	1	MGLPRGPLASLLLVCMVLCQAASPCRAVFEAEVTLKAGGAEQPGQALGVFMGCPG	60
QY	61	QEPALFSTNDNDDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG	120
Db	61	QEPALFSTNDNDDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG	120
QY	121	KGPPFORLNQKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREEIAK	180
Db	121	KGPPFORLNQKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREEIAK	180
QY	181	YELFGHVAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLGVLPGTSMVQVAT	240
Db	181	YELFGHVAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLGVLPGTSMVQVAT	240
QY	241	DEDDAIYTYNGVVAYSIHQSEPKDPHDLMTFTHRSTGTISVSSGLDREKVPYETLTIOA	300
Db	241	DEDDAIYTYNGVVAYSIHQSEPKDPHDLMTFTHRSTGTISVSSGLDREKVPYETLTIOA	300
QY	301	TMMDGSGSTTTAVAVVEILDANDNAPMFPDQKYEAHVPENAVGVHVEVQRLTVDLDPNSP	360
Db	301	TMMDGSGSTTTAVAVVEILDANDNAPMFPDQKYEAHVPENAVGVHVEVQRLTVDLDPNSP	360
QY	361	AWRATYLMGGDDGDHFTTTHPESNOGILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420
Db	361	AWRATYLMGGDDGDHFTTTHPESNOGILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420

XX 31-MAY-2001; 2001US-0294225P.
XX (CHIR) CHIRON CORP.
XX
XX Reinhard C, Klinger J, Jefferson AB, Escobedo J, Randazo F;
XX Winter J, Goodson R;
XX WPI; 2003-140501/13.
XX N-PSDB; ABZ24736.
XX
XX Inhibiting migration and proliferation of P-cadherin expressing cancer
XX for treating cancer, especially digestive cancer, characterized by
XX overexpression of P-cadherin, involves administering a P-cadherin
XX antagonist.
XX
XX Disclosure; Page 10-11; 129pp; English.
XX
XX The present sequence is the protein sequence for human P-cadherin
XX (placental cadherin), a calcium-dependent cellular adhesion protein. The
XX invention provides methods of treating or diagnosing cancers involving P-
XX cadherin expression using ligands that target P-cadherin, especially
XX human anti-P-cadherin antibodies. A claimed method of treating a cancer
XX characterised by the overexpression and/or upregulation of P-cadherin
XX comprises the administration of a P-cadherin antagonist, optionally
XX conjugated to a therapeutic agent. The migration, adhesion and/or
XX proliferation of the cancer is inhibited, and the method is especially
XX useful for treating or preventing a digestive cancer such as colon or
XX colorectal cancer. The antagonists may be an anti-P-cadherin antibody or
XX its fragment, a ribozyme or antisense oligonucleotide. A transgenic
XX animal that expresses a recombinant (human) antibody that specifically
XX binds P-cadherin is also claimed. A method of determining the presence of
XX cancer involves determining the expression level of P-cadherin in a cell
XX sample. Also provided are screens for identifying anti-P-cadherin
XX antibodies have therapeutic activity
XX
XX Sequence 829 AA;

Query Match 100.0%; Score 4369; DB 6; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLPRGPLASLLQLQVCWLOCAAEPCRAVFAEAVTLEAGGAEQEPQALGKVFMCPCG 60
DB 1 MGLPRGPLASLLQLQVCWLOCAAEPCRAVFAEAVTLEAGGAEQEPQALGKVFMCPCG 60
QY 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG 120
DB 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG 120
QY 121 KGPPFQRLNQLSKNDRDKTFYSGITGPGADSPPEGVFAVEKETGWLNNKPLDREBIAK 180
DB 121 KGPPFQRLNQLSKNDRDKTFYSGITGPGADSPPEGVFAVEKETGWLNNKPLDREBIAK 180
QY 181 YELFGHAVSENGASVEDPMNISITVTDQNDHKPKFTQDTPRGSLVGLPGTSVMQVAT 240
DB 181 YELFGHAVSENGASVEDPMNISITVTDQNDHKPKFTQDTPRGSLVGLPGTSVMQVAT 240
QY 241 DEDDAITYNGVAVSIHQSPKDPKPHDLMTFTIHRSTGTISVISGLDREKVPYTLTQA 300
DB 241 DEDDAITYNGVAVSIHQSPKDPKPHDLMTFTIHRSTGTISVISGLDREKVPYTLTQA 300
QY 301 TDMGDGSTTAVAVVEILDANDANAPMDPKYEAHVPENAVGHVEQRLTVTDLDAPNSP 360
DB 301 TDMGDGSTTAVAVVEILDANDANAPMDPKYEAHVPENAVGHVEQRLTVTDLDAPNSP 360
QY 361 AWRATYLLMGDDGDHFTTTHPSNQGLITTRKGLDPEAKNQHTLVVEVNEAPFVLKL 420
DB 361 AWRATYLLMGDDGDHFTTTHPSNQGLITTRKGLDPEAKNQHTLVVEVNEAPFVLKL 420
QY 421 PTSTATIYVHVVEDVNEAPFVPPSKVVEVQEGIPTGEVPCVYTAEDDPKQKISYRILR 480
DB 421 PTSTATIYVHVVEDVNEAPFVPPSKVVEVQEGIPTGEVPCVYTAEDDPKQKISYRILR 480

QY 481 DPAGWLAMPDSDSQVTAAGTLDREDSQFVRNNIYEVWLAMDNGSPPTTGTGTLTLLID 540
DB 481 DPAGWLAMPDSDSQVTAAGTLDREDSQFVRNNIYEVWLAMDNGSPPTTGTGTLTLLID 540
QY 541 VNDHGVPPEPRQITICNQSPVRHVLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD 600
DB 541 VNDHGVPPEPRQITICNQSPVRHVLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD 600
QY 601 TVVLSLKKFLKQDYDVHLSLDHGNKEQLTVIRATVCDCHGHVETCPGPKGGFLLPVL 660
DB 601 TVVLSLKKFLKQDYDVHLSLDHGNKEQLTVIRATVCDCHGHVETCPGPKGGFLLPVL 660
QY 661 GAVLALLFLLVLLVLLVRKKIKPELLLPEDDTRDNVFFYEGEGEEDQDYDITQLHR 720
DB 661 GAVLALLFLLVLLVLLVRKKIKPELLLPEDDTRDNVFFYEGEGEEDQDYDITQLHR 720
QY 721 GLEARPEVVLNDVAPTIIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
DB 721 GLEARPEVVLNDVAPTIIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
QY 781 DYEGSSDAASLSLSSASDQDDYDYLNEWGSRPKKLADMYGGGEDD 829
DB 781 DYEGSSDAASLSLSSASDQDDYDYLNEWGSRPKKLADMYGGGEDD 829
RESULT 7
ADCL5497
ID ADCL5497 standard; protein; 829 AA.
XX
XX AC ADCL5497;
XX
XX AC ADCL5497;
DT 18-DEC-2003 (first entry)
XX
XX DE Human basal cell marker Cadherin-3.
XX
XX KW Human; basal cell marker; cadherin-3; tumour; P-cadherin;
XX KW matrix metalloproteinase 14;
XX KW cadherin EGF LAG seven-pass G-type receptor 2;
XX KW EGF-like domain Multiple 2; mammary lactation gland; breast tumour;
XX KW cancer; cytostatic.
XX
XX OS Homo sapiens.
XX
XX PN US2003086934-A1.
XX
XX PD 08-MAY-2003.
XX
XX PF 26-JUL-2001; 2001US-00916849.
XX
XX PR 26-JUL-2000; 2000US-0220967P.
XX
XX PA (BOTS/) BOTSTEIN D.
XX PA (BROW/) BROWN P O.
XX PA (PERO/) PEROU C M.
XX PA (RING/) RING C M.
XX PA (ROSS/) ROSS D.
XX PA (SEIT/) SEITZ R.
XX PA (VRIJ/) VAN DE RIJN J M.
XX
XX PI Botstein D, Brown PO, Perou CM, Ring B, Ross D, Seitz R;
XX Van De Rijn JM;
XX
XX WPI; 2003-786931/74.
XX
XX DR
XX PT Classifying breast tumor by detecting expression or activity of gene
XX encoding cadherin 3, matrix metalloproteinase 14, cadherin EGF LAG seven-
XX pass G-type receptor 2 in tumor sample.
XX
XX PS Claim 1; SEQ ID NO 1; 88pp; English.
XX
XX CC The invention relates to classifying (M1) a tumour comprising detecting
XX expression or activity of a gene encoding cadherin 3 or P-cadherin
CC

CC ,matrix metalloproteinase 14 , or cadherin EGF LAG seven-pass G-type
CC receptor 2 or EGF-like domain, Multiple 2 (PI-3) or at least two genes
CC chosen from gene encoding (PI), (P2) or (P3) in the tumour sample, and
CC classifying tumour as belonging to a tumour subclass based on results of
CC detecting step. Also included are testing a subpredictor for expression of P1-
CC P3 and providing diagnostic, prognostic, or predictive information based
CC on the detecting step, stratifying the subject for a clinical trial based
CC on the detecting step, or selecting a treatment based on the detecting
CC step, an antibody that specifically binds to an epitope found in (P1),
CC (P2) or (P3) (where the antibody recognises basal cells in normal mammary
CC lactation glands), methods of testing a compound or a combination of
CC compounds for activity against tumours. The method (M1) is useful for
CC classifying a tumour (breast tumour), as belonging to a tumour subclass
CC such as basal tumour subclass. The antibody is useful for treating a
CC subject in need of treatment for cancer. The tumour is a breast tumour,
CC and where the method further involves identifying the tumour as belonging
CC to a basal tumour subclass. The antibody is conjugated with a toxic
CC molecule. The present sequence is human Cadherin-3.
XX
SQ Sequence 829 AA;
Query Match 100.0%; Score 4369; DB 7; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLPRGLASLLQLQVLCQAASEPCRAVFREAETLEAGGAEQPGQALGKVFMGCPG 60
Db 1 MGLPRGLASLLQLQVLCQAASEPCRAVFREAETLEAGGAEQPGQALGKVFMGCPG 60
QY 61 QEPALFSTNDNDFTVRNGETVQRRSLKERNPLKIPPSKRILRRHXRDMVWVAPISVPENG 120
Db 61 QEPALFSTNDNDFTVRNGETVQRRSLKERNPLKIPPSKRILRRHXRDMVWVAPISVPENG 120
QY 121 KGPFPORLNQKSKNDRTKIFYSITGPGADSPGEGVFAVEKETGMLLNKPLDREIAK 180
Db 121 KGPFPORLNQKSKNDRTKIFYSITGPGADSPGEGVFAVEKETGMLLNKPLDREIAK 180
QY 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHDKPKFTQDTRFGSVLBSGLVPGTSMQVAT 240
Db 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHDKPKFTQDTRFGSVLBSGLVPGTSMQVAT 240
QY 241 DEDDAITYNGVAVYHSGQPKDHDLMFTTHRSSTGTSVSSGLDRKPEYTLTIOA 300
Db 241 DEDDAITYNGVAVYHSGQPKDHDLMFTTHRSSTGTSVSSGLDRKPEYTLTIOA 300
QY 301 TMDGPGSTTTAVAVVEILDANDNAPMDPKYEAHVPENAVGVORLTVTDLDPNSP 360
Db 301 TMDGPGSTTTAVAVVEILDANDNAPMDPKYEAHVPENAVGVORLTVTDLDPNSP 360
QY 361 AWRATYLMGGDDGDHFTTTHPESNQGLITRKGLDFAKQHTLYVEVTNEAPFVLKL 420
Db 361 AWRATYLMGGDDGDHFTTTHPESNQGLITRKGLDFAKQHTLYVEVTNEAPFVLKL 420
QY 421 PTSTATIVHVEDVNEAPVPPSKVVEQEGIPTGEPVCVYTAEPDPEKNIKISRIIR 480
Db 421 PTSTATIVHVEDVNEAPVPPSKVVEQEGIPTGEPVCVYTAEPDPEKNIKISRIIR 480
QY 481 DPAGWLAMPDSCGVATGCTLDREDSQFVRNNIYEVWVLAMDNQSPPTTGTGLTLTLD 540
Db 481 DPAGWLAMPDSCGVATGCTLDREDSQFVRNNIYEVWVLAMDNQSPPTTGTGLTLTLD 540
QY 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
QY 601 TVVLSLKFKLQDPTVDHLSLSHGNKEQLTVIRATVCDCHGVETCPGPKWGGFLLPVL 660
Db 601 TVVLSLKFKLQDPTVDHLSLSHGNKEQLTVIRATVCDCHGVETCPGPKWGGFLLPVL 660
QY 661 GAVLALLFLLVLLVLLVLRKKIKKEPLLPEDDTRDNVIFYEGEGGEEDQYDITQLHR 720
Db 661 GAVLALLFLLVLLVLLVLRKKIKKEPLLPEDDTRDNVIFYEGEGGEEDQYDITQLHR 720

QY 721 GLEARPEVVLNRDVAPTIPTPMYPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVVLNRDVAPTIPTPMYPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
QY 781 DYEGSGDAASLSLTSASDQDDYDVLNWSRFRKFLADMYGGEDD 829
Db 781 DYEGSGDAASLSLTSASDQDDYDVLNWSRFRKFLADMYGGEDD 829
RESULT 8
ADD14190
ID ADD14190 standard; protein; 829 AA.
XX
AC ADD14190;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human src biomarker polypeptide SEQ ID NO:379.
XX
KW predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human.
XX
OS Homo sapiens.
XX
PN WO2003062395-A2.
XX
PD 31-JUL-2003.
XX
PF 17-JAN-2003; 2003WO-US001981.
XX
PR 18-JAN-2002; 2002US-0350061P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Huang F, Fairchild CR, Lee FY, Shaw P;
XX
WPI; 2003-636735/60.
DR N-PSDB; ADD14795.
XX
PT New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
XX tyrosine kinase pathways.
PS Claim 10; SEQ ID NO 379; 139pp; English.
XX
CC The present invention describes a predictor set comprising a plurality of
CC polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.

QY 1 MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLFAGGAEPGQALGKVFMGCPG 60
DB 1 MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLFAGGAEPGQALGKVFMGCPG 60
QY 61 QBPALFSTNDNDFTVRNGETVQERRSLKERNPLKIPPSKRIILRRHKRDWVAPISVPENG 120
DB 61 QBPALFSTNDNDFTVRNGETVQERRSLKERNPLKIPPSKRIILRRHKRDWVAPISVPENG 120
QY 121 KGPFFQRLNQLKSNKDRDKTFIYSITGPGADSPPEGFAVEKETGWLILNKLPLDREEIAK 180
DB 121 KGPFFQRLNQLKSNKDRDKTFIYSITGPGADSPPEGFAVEKETGWLILNKLPLDREEIAK 180
QY 181 YELFGHAVSENGASVEDPNNISIIIVTDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVTTAT 240
DB 181 YELFGHAVSENGASVEDPNNISIIIVTDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVTTAT 240
QY 241 DEDDALTYTNGVAVYSIHQSEPKDHDLMFTTHRSSTGTSIVSSGLDREKPEYTLTIOA 300
DB 241 DEDDALTYTNGVAVYSIHQSEPKDHDLMFTTHRSSTGTSIVSSGLDREKPEYTLTIOA 300
QY 301 TMDGSGSTTTAVAVVEILDANDNAMPDPQKYEAHVPENAVGHEVQRLTVTDLDAENSP 360
DB 301 TMDGSGSTTTAVAVVEILDANDNAMPDPQKYEAHVPENAVGHEVQRLTVTDLDAENSP 360
QY 361 ANRATYILMGDDGDHFTTTHPESNQGLITRKGDLDFEAKNQHTLYVEVTNEAPFVLKL 420
DB 361 ANRATYILMGDDGDHFTTTHPESNQGLITRKGDLDFEAKNQHTLYVEVTNEAPFVLKL 420
QY 421 PTSTATIVVHVEDVNEAPFVPPSKVVEQEGTPTGEPVCVVTAEPPDKENKISRIILR 480
DB 421 PTSTATIVVHVEDVNEAPFVPPSKVVEQEGTPTGEPVCVVTAEPPDKENKISRIILR 480
QY 481 DPAGWLAMPDQSGQVAVGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGLTLTLID 540
DB 481 DPAGWLAMPDQSGQVAVGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGLTLTLID 540
QY 541 VNDHGVPPEPRQITICNQSPVRHVLNITDKOLSPHTSPFQAQITDSDIYWTAEVNEED 600
DB 541 VNDHGVPPEPRQITICNQSPVRHVLNITDKOLSPHTSPFQAQITDSDIYWTAEVNEED 600
QY 601 TVVLSLKKFLKQDTYDVHLSLDHGNKEQLTVIRATVCDCHGVETCPGWPFGGFLPVL 660
DB 601 TVVLSLKKFLKQDTYDVHLSLDHGNKEQLTVIRATVCDCHGVETCPGWPFGGFLPVL 660
QY 661 GAVLALLFLLVLLVLRKKRIKEPLLPEDDTRDNVYFYGEGGEDQDYDITQLHR 720
DB 661 GAVLALLFLLVLLVLRKKRIKEPLLPEDDTRDNVYFYGEGGEDQDYDITQLHR 720
QY 721 GLEARPEVVLNDVAPTIITPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
DB 721 GLEARPEVVLNDVAPTIITPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
QY 781 DYEGSGSDAASLSLTSASDQDDYDYLNEGSRFKKLADMYGGGEDD 829
DB 781 DYEGSGSDAASLSLTSASDQDDYDYLNEGSRFKKLADMYGGGEDD 829
RESULT 10
ADN39578
ID ADN39578 standard; protein; 829 AA.
XX
AC ADN39578;
DT
XT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A178.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;

KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine.
OS Homo sapiens.
XX WO2003042661-A2.
XX 22-MAY-2003.
XX 13-NOV-2002; 2002WO-US036810.
XX 13-NOV-2001; 2001US-0350666P.
XX 21-NOV-2001; 2001US-0332464P.
XX 29-NOV-2001; 2001US-0334393P.
XX 03-DEC-2001; 2001US-0335394P.
XX 14-DEC-2001; 2001US-0340376P.
XX 08-JAN-2002; 2002US-0347211P.
XX 10-JAN-2002; 2002US-0347349P.
XX 08-FEB-2002; 2002US-0355250P.
XX 13-FEB-2002; 2002US-0356714P.
XX 20-FEB-2002; 2002US-0359077P.
XX 29-MAR-2002; 2002US-0368099P.
XX 04-APR-2002; 2002US-0370110P.
XX 12-APR-2002; 2002US-0372246P.
XX 05-JUN-2002; 2002US-0386614P.
XX 16-JUL-2002; 2002US-0396839P.
XX 22-JUL-2002; 2002US-039775P.
XX 22-JUL-2002; 2002US-0397845P.
XX 09-SEP-2002; 2002US-0409450P.
XX (EOB-) EOS BIOTECHNOLOGY INC.
XX PA Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
XX PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX DR WPI: 2003-468649/44.
XX N-PSDB; ADN39577.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX Claim 12; SEQ ID NO A178; 1385pp; English.
XX The invention relates to nucleic acids and proteins (ADN39683-ADN40064)
XX whose expression is upregulated or downregulated in specific cancers or
XX other diseases such as angiogenic or fibrotic disorders, and to methods
XX of determining the presence or absence of a pathological cell in a
XX patient by detecting a nucleic acid at least 80% identical to those of
XX the invention or by detecting a polypeptide of the invention. The
XX invention also relates to expression vectors and host cells comprising a
XX nucleic acid of the invention; antibodies which specifically bind a
XX polypeptide of the invention; use of such antibodies for drug targeting;
XX and methods of screening for modulators of activity or expression of the
XX polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX antibodies and methods are useful for diagnosing, prognosing and treating
XX cancer and other conditions such as psoriasis, ischaemia, heart disease,
XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX neovascularisation syndromes, scarring and uterine fibroids. They may
XX also be useful in wound healing and in contraception. The present
XX sequence represents a polypeptide of the invention.
XX
XX Sequence 829 AA;
XX Query Match 100.0%; Score 4369; DB 7; Length 829;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLFAGGAEPGQALGKVFMGCPG 60
DB 1 MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLFAGGAEPGQALGKVFMGCPG 60
QY 61 QBPALFSTNDNDFTVRNGETVQERRSLKERNPLKIPPSKRIILRRHKRDWVAPISVPENG 120

Db 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKILRRHKRDWVAPISVPENG 120
Qy 121 KGPPQRLNOLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLLNKPDLREETAK 180
Db 121 KGPPQRLNOLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLLNKPDLREETAK 180
Qy 181 YELFGHAVSENGASVEDPMMISIIIVTDNDHKPKFTQDTPRGSVLEGLVPGTSMQVYAT 240
Db 181 YELFGHAVSENGASVEDPMMISIIIVTDNDHKPKFTQDTPRGSVLEGLVPGTSMQVYAT 240
Qy 241 DEDDAIYTYNGWVAYSIIHQEPKDPHLMFTIHRSTGTISVSSGLDREKVPEYTLTQA 300
Db 241 DEDDAIYTYNGWVAYSIIHQEPKDPHLMFTIHRSTGTISVSSGLDREKVPEYTLTQA 300
Qy 301 TDMGDGSGTTTAVAVVEILDANDNAPFDPQKYEHAHPVENAVGHEVORLTVTDLDAENSP 360
Db 301 TDMGDGSGTTTAVAVVEILDANDNAPFDPQKYEHAHPVENAVGHEVORLTVTDLDAENSP 360
Qy 361 AWRATYILMGDDGDHFTITTHPSNQILITTRKGLDFEAKNOHTLVVEVTNEAPFVLKL 420
Db 361 AWRATYILMGDDGDHFTITTHPSNQILITTRKGLDFEAKNOHTLVVEVTNEAPFVLKL 420
Qy 421 PTSTATIVHVEDVNEAPVFVPPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480
Db 421 PTSTATIVHVEDVNEAPVFVPPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480
Qy 481 DPAGWLAMPDSGGQVTAAGTLDREDEQFVRNNIYEVMLANDNGSPPTGTGTLTLLID 540
Db 481 DPAGWLAMPDSGGQVTAAGTLDREDEQFVRNNIYEVMLANDNGSPPTGTGTLTLLID 540
Qy 541 VNDHGVPPEPQIITICNQSPVRHVLNITDKLSPTGPFQAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGVPPEPQIITICNQSPVRHVLNITDKLSPTGPFQAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSLKKFLKQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGFIPLVL 660
Db 601 TVVLSLKKFLKQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGFIPLVL 660
Qy 661 GAVLALLFLLAVLALLVLRKKIKIPELLIPEDTDRDNVYVYGGEGGEEDQDITOLHR 720
Db 661 GAVLALLFLLAVLALLVLRKKIKIPELLIPEDTDRDNVYVYGGEGGEEDQDITOLHR 720
Qy 721 GLEARPEVLNDVAPTIPTPMYRPRANPDEIGNFTIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVLNDVAPTIPTPMYRPRANPDEIGNFTIENLKAANTDPTAPPYDTLLVF 780
Qy 781 DYEGSGSDAASLSLTSSASDQDQDYDYLNEWGSRFKKLADMYCGGEDD 829
Db 781 DYEGSGSDAASLSLTSSASDQDQDYDYLNEWGSRFKKLADMYCGGEDD 829

RESULT 11
ADN39545
ID ADN39545 standard; protein; 829 AA.

XX AC ADN39545;

XX DT 17-JUN-2004 (first entry)

XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A145.

XX KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine.

XX OS Homo sapiens.

XX PN WO2003042661-A2.

XX PD 22-MAY-2003 .
XX PF 13-NOV-2002; 2002WO-US036810 .
XX PR 13-NOV-2001; 2001US-0350666P .
PR 21-NOV-2001; 2001US-0332464P .
PR 29-NOV-2001; 2001US-0334393P .
PR 03-DEC-2001; 2001US-0335394P .
PR 14-DEC-2001; 2001US-0340376P .
PR 08-JAN-2002; 2002US-0347211P .
PR 10-JAN-2002; 2002US-0347349P .
PR 08-FEB-2002; 2002US-035250P .
PR 13-FEB-2002; 2002US-0356714P .
PR 20-FEB-2002; 2002US-0359077P .
PR 29-MAR-2002; 2002US-0368809P .
PR 04-APR-2002; 2002US-0370110P .
PR 12-APR-2002; 2002US-0372246P .
PR 05-JUN-2002; 2002US-0386614P .
PR 16-JUL-2002; 2002US-0396839P .
PR 22-JUL-2002; 2002US-0397775P .
PR 22-JUL-2002; 2002US-0397845P .
PR 09-SEP-2002; 2002US-0409450P .
XX PA (SO5B-) EOS BIOTECHNOLOGY INC.
XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilton KE, Zlotnik A;
XX DR WPI; 2003-468649/44 .
DR N-PSDB; ADN39544 .
XX PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample .
XX PS Claim 12; SEQ ID NO A145; 1385pp; English.
XX CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acid, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX SQ Sequence 829 AA;

Query Match 100.0%; Score 4369; DB 7; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGLPRGPLASILLQVCWLCQAASEPCRAVFREAEVTLAAGGAEQEPQALGKVFMCPCG 60
Db 1 MGLPRGPLASILLQVCWLCQAASEPCRAVFREAEVTLAAGGAEQEPQALGKVFMCPCG 60
Qy 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKILRRHKRDWVAPISVPENG 120
Qy 121 KGPPQRLNOLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLLNKPDLREETAK 180
Db 121 KGPPQRLNOLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLLNKPDLREETAK 180

QY	181	YELFGHAVSENGASVEDPMNISIIVTDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVTTAT	240
Db	181	YELFGHAVSENGASVEDPMNISIIVTDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVTTAT	240
QY	241	DEDDAIYTVNGVAVSIHSGQPKDPHDLMTFTHRSSTGTSVSSGDLREKVPYETLTIOA	300
Db	241	DEDDAIYTVNGVAVSIHSGQPKDPHDLMTFTHRSSTGTSVSSGDLREKVPYETLTIOA	300
QY	301	TMDGSGSTTTAVAVVEILDANDNPMFDPQKYEAHVPENAVHGVQRLTVTDLDAVNSP	360
Db	301	TMDGSGSTTTAVAVVEILDANDNPMFDPQKYEAHVPENAVHGVQRLTVTDLDAVNSP	360
QY	361	AWRATYLINGDDGDHFTTTHPESQGLITTRKGLDFAKQHTLIYEVVTEAPFVLKL	420
Db	361	AWRATYLINGDDGDHFTTTHPESQGLITTRKGLDFAKQHTLIYEVVTEAPFVLKL	420
QY	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPGCEPVVTVTAEDPKENOKISVRIILR	480
Db	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPGCEPVVTVTAEDPKENOKISVRIILR	480
QY	481	DPAGWLAMPDQSGQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGTLTLLID	540
Db	481	DPAGWLAMPDQSGQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGTLTLLID	540
QY	541	VNDHGVPPEPRQITICNQSPVRHVLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD	600
Db	541	VNDHGVPPEPRQITICNQSPVRHVLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD	600
QY	601	TVVLSLKKFLKQDVTYVHLSLDHGNKEQLTVIRATVCDCHGVETCPGPKGGFLPVL	660
Db	601	TVVLSLKKFLKQDVTYVHLSLDHGNKEQLTVIRATVCDCHGVETCPGPKGGFLPVL	660
QY	661	GAVLALLFLLVLLVLRKKRIKEPLLPEDDTRDNVFFYEGEGEDQDYDITQLHR	720
Db	661	GAVLALLFLLVLLVLRKKRIKEPLLPEDDTRDNVFFYEGEGEDQDYDITQLHR	720
QY	721	GLEARPEVLRNDVAPTIPTMYRPRPANPBEIGNFIENLKAANTDPTAPPYDILLVF	780
Db	721	GLEARPEVLRNDVAPTIPTMYRPRPANPBEIGNFIENLKAANTDPTAPPYDILLVF	780
QY	781	DYEGSGSDAASLSLTSASDQDDQDYDLNWSRPFKLLADMYGGGEDD	829
Db	781	DYEGSGSDAASLSLTSASDQDDQDYDLNWSRPFKLLADMYGGGEDD	829
RESULT 12			
ID	ADN39465	standard; protein; 829 AA.	
AC	ADN39465;		
XX	ADN39465;		
DT	17-JUN-2004	(first entry)	
XX	Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A65.		
DE	Human;		
XX	Human; differential expression; cancer; angiogenic disorder;		
KW	fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;		
KW	inflammatory disease; autoimmune disease;		
KW	retinal neovascularisation syndrome; scarring; uterine fibroid;		
KW	detection; diagnosis; prognosis; drug screening; drug targeting;		
KW	wound healing; contraception; cytostatic; cardiac; immunomodulatory;		
KW	vulnerary; gene therapy; vaccine.		
XX	Homo sapiens.		
OS			
XX	WO2003042661-A2.		
PN			
XX	22-MAY-2003.		
XX			
PD	13-NOV-2002; 2002WO-US036810.		
XX			
PF			
XX	13-NOV-2001; 2001US-0350666P.		

PR	21-NOV-2001; 2001US-0332464P.		
PR	29-NOV-2001; 2001US-0334393P.		
PR	03-DEC-2001; 2001US-0335394P.		
PR	14-DEC-2001; 2001US-0340376P.		
PR	08-JAN-2002; 2002US-0347211P.		
PR	10-JAN-2002; 2002US-0347349P.		
PR	08-FEB-2002; 2002US-035250P.		
PR	13-FEB-2002; 2002US-0356714P.		
PR	20-FEB-2002; 2002US-0359077P.		
PR	29-MAR-2002; 2002US-036809P.		
PR	04-APR-2002; 2002US-0370110P.		
PR	12-APR-2002; 2002US-0372246P.		
PR	05-JUN-2002; 2002US-0386614P.		
PR	16-JUL-2002; 2002US-0396839P.		
PR	22-JUL-2002; 2002US-039775P.		
PR	22-JUL-2002; 2002US-0397845P.		
PR	09-SEP-2002; 2002US-0409450P.		
XX	(EOSB-) EOS BIOTECHNOLOGY INC.		
PA			
XX			
PI	Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;		
PI	Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;		
XX			
XX			
DR	WPI: 2003-468649/44.		
DR	N-PSDB; ADN39464.		
XX			
PT	Determining the presence or absence of a pathological cell in a patient,		
PT	useful for diagnosing, prognosing or treating cancer, comprises detecting		
PT	a nucleic acid in a biological sample.		
XX			
PS	Claim 12; SEQ ID NO A65; 1385pp; English.		
CC	The invention relates to nucleic acids and proteins (ADN39683-ADN40064)		
CC	whose expression is upregulated or downregulated in specific cancers or		
CC	other diseases such as angiogenic or fibrotic disorders, and to methods		
CC	of determining the presence or absence of a pathological cell in a		
CC	patient by detecting a nucleic acid at least 80% identical to those of		
CC	the invention or by detecting a polypeptide of the invention. The		
CC	invention also relates to expression vectors and host cells comprising a		
CC	nucleic acid of the invention; antibodies which specifically bind a		
CC	polypeptide of the invention; use of such antibodies for drug targeting;		
CC	and methods of screening for modulators of activity or expression of the		
CC	polypeptides and nucleic acids. The nucleic acids, polypeptides, and		
CC	antibodies and methods are useful for diagnosing, prognosing and treating		
CC	cancer and other conditions such as psoriasis, ischaemia, heart disease,		
CC	atherosclerosis, inflammatory diseases, autoimmune diseases, retinal		
CC	neovascularisation syndromes, scarring and uterine fibroids. They may		
CC	also be useful in wound healing and in contraception. The present		
CC	sequence represents a polypeptide of the invention.		
XX			
SQ	Sequence 829 AA;		
Query Match			
Best Local Similarity 100.0%; Score 4369; DB 7; Length 829;			
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	MGLPRGLASLLQLQVCWLQCAASPCRAVFEAEVTLAAGGAEQPGQALGKVFMGCPG	60
Db	1	MGLPRGLASLLQLQVCWLQCAASPCRAVFEAEVTLAAGGAEQPGQALGKVFMGCPG	60
QY	61	QBPALFSTNDQDFTVRNGETVOERSLKERNPLKIPPSKRIILRRHKRDVVVAPISVPENG	120
Db	61	QBPALFSTNDQDFTVRNGETVOERSLKERNPLKIPPSKRIILRRHKRDVVVAPISVPENG	120
QY	121	KGPPQRLNOLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKLKPDREEIAK	180
Db	121	KGPPQRLNOLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKLKPDREEIAK	180
QY	181	YELFGHAVSENGASVEDPMNISIIVTDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVTTAT	240
Db	181	YELFGHAVSENGASVEDPMNISIIVTDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVTTAT	240
QY	241	DEDDAIYTVNGVAVSIHSGQPKDPHDLMTFTHRSSTGTSVSSGDLREKVPYETLTIOA	300

Db 241 DEDDAIYTVNGVWVAYSIHSQEPKDPHDLMTFIHRSTGTISVISGLDREKVPVETLTIQA 300
Qy 301 TDMGDGSGTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVEQRLTVTDLDAENSP 360
Db 301 TDMGDGSGTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVEQRLTVTDLDAENSP 360
Qy 361 AWRATYILMGDDGDHFTIITHPESNQGILTRKGLDPEAKNQHTLVVEVTNEAPFVLKL 420
Db 361 AWRATYILMGDDGDHFTIITHPESNQGILTRKGLDPEAKNQHTLVVEVTNEAPFVLKL 420
Qy 421 PTSTATIVVHVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDPKENQKISYRILR 480
Db 421 PTSTATIVVHVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDPKENQKISYRILR 480
Qy 481 DPAGWLAMPDPSGQVTVAGTLDREDEQFVRNNIYEVNVLAMNDGSPPTGTGTLTLLID 540
Db 481 DPAGWLAMPDPSGQVTVAGTLDREDEQFVRNNIYEVNVLAMNDGSPPTGTGTLTLLID 540
Qy 541 VNDHGPVPEPQIITCNQSPVRHVNLITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPQIITCNQSPVRHVNLITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSLKKFLKQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Db 601 TVVLSLKKFLKQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Qy 661 GAVLALLFLLLVLLLVKRRKIKIPEPLLLPDDTRDNVFFYGGEGGEEDQDITQLHR 720
Db 661 GAVLALLFLLLVLLLVKRRKIKIPEPLLLPDDTRDNVFFYGGEGGEEDQDITQLHR 720
Qy 721 GLEARPEVLNDVAPTIPTPMYRPRPNDPEIGNFIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVLNDVAPTIPTPMYRPRPNDPEIGNFIENLKAANTDPTAPPYDTLLVF 780
Qy 781 DYEAGSGDAASLSSITSSASDQDQDYDYLNEWGSRFKKLADMYGGGDD 829
Db 781 DYEAGSGDAASLSSITSSASDQDQDYDYLNEWGSRFKKLADMYGGGDD 829
RESULT 13
ADL06565
ID ADL06565 standard; protein; 829 AA.
XX AC ADL06565;
XX DT 20-MAY-2004 (first entry)
XX DE Human tumour-associated antigenic target (TAT) polypeptide #64.
XX KW Human; tumour-associated antigenic target; TAT; cell death; tumour;
XX KW cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO2004016225-A2.
XX XX 26-FEB-2004.
XX FD 19-AUG-2003; 2003WO-US025892.
XX PR 19-AUG-2002; 2002US-0404809P.
XX PR 21-AUG-2002; 2002US-0405645P.
XX PR 23-SEP-2002; 2002US-0413192P.
XX PR 15-OCT-2002; 2002US-0419008P.
XX PR 15-NOV-2002; 2002US-0426847P.
XX PR 02-JUL-2003; 2003US-0484959P.
XX PA (GETH) GENENTECH INC.
XX PI Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;
PI Spencer SD, Wu TD, Zhang Z;
XX

DR WPI; 2004-257144/24.
DR N-PSDB; ADL06487.
XX New antibody that binds to a tumor-associated antigenic target (TAT) polypeptide, useful for preparing a composition for diagnosing or treating cancer.
PT PT
PT PT
PT PT
PS Claim 2; SEQ ID NO 145; 319pp; English.
XX
CC The present invention relates to the isolation of human tumour-associated antigenic target (TAT) polynucleotide and polypeptide sequences. Also disclosed is an antibody that binds to a TAT polypeptide. The antibody is a monoclonal antibody, an antibody fragment, a chimeric antibody or a humanised antibody. It is conjugated to a growth inhibitory agent. It is produced in bacteria or in CHO cells and induces death of a cell to which it binds. The antibody is useful for preparing a composition for diagnosing or treating tumours and cancer. The present sequence CC represents a human TAT polypeptide of the invention.
XX SQ Sequence 829 AA;
Query Match 100.0%; Score 4369; DB 8; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGLPRGPLASLLQLQVCWLOCAASEPCRAVFREAEVTLAEAGAEQEPQALGKVFMGCPG 60
Db 1 MGLPRGPLASLLQLQVCWLOCAASEPCRAVFREAEVTLAEAGAEQEPQALGKVFMGCPG 60
Qy 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIPSKILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIPSKILRRHKRDWVAPISVPENG 120
Qy 121 KGFPQRLNQLKSKORDTKIFYISITGPGADSPGEGVAFVEKETGWLILNKPLDREELAK 180
Db 121 KGFPQRLNQLKSKORDTKIFYISITGPGADSPGEGVAFVEKETGWLILNKPLDREELAK 180
Qy 181 YELFGHAVSENGASVEDPMNISIIIVTDNDHKPKFTQDTFRGSLGVLGPGTSMQVYAT 240
Db 181 YELFGHAVSENGASVEDPMNISIIIVTDNDHKPKFTQDTFRGSLGVLGPGTSMQVYAT 240
Qy 241 DEDDAIYTVNGVWVAYSIHSQEPKDPHDLMTFIHRSTGTISVISGLDREKVPVETLTIQA 300
Db 241 DEDDAIYTVNGVWVAYSIHSQEPKDPHDLMTFIHRSTGTISVISGLDREKVPVETLTIQA 300
Qy 301 TDMGDGSGTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVEQRLTVTDLDAENSP 360
Db 301 TDMGDGSGTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVEQRLTVTDLDAENSP 360
Qy 361 AWRATYILMGDDGDHFTIITHPESNQGILTRKGLDPEAKNQHTLVVEVTNEAPFVLKL 420
Db 361 AWRATYILMGDDGDHFTIITHPESNQGILTRKGLDPEAKNQHTLVVEVTNEAPFVLKL 420
Qy 421 PTSTATIVVHVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDPKENQKISYRILR 480
Db 421 PTSTATIVVHVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDPKENQKISYRILR 480
Qy 481 DPAGWLAMPDPSGQVTVAGTLDREDEQFVRNNIYEVNVLAMNDGSPPTGTGTLTLLID 540
Db 481 DPAGWLAMPDPSGQVTVAGTLDREDEQFVRNNIYEVNVLAMNDGSPPTGTGTLTLLID 540
Qy 541 VNDHGPVPEPQIITCNQSPVRHVNLITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPQIITCNQSPVRHVNLITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSLKKFLKQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Db 601 TVVLSLKKFLKQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Qy 661 GAVLALLFLLLVLLLVKRRKIKIPEPLLLPDDTRDNVFFYGGEGGEEDQDITQLHR 720
Db 661 GAVLALLFLLLVLLLVKRRKIKIPEPLLLPDDTRDNVFFYGGEGGEEDQDITQLHR 720

```
QY 721 GLEARPEVLRNDVAPTIPTMYRPRRPNPDEIGNFIENLKAANTDPTAPPYDILLVF 780
Db 721 GLEARPEVLRNDVAPTIPTMYRPRRPNPDEIGNFIENLKAANTDPTAPPYDILLVF 780
QY 781 DYEGSGDAASLSLTSASDQDDYDYLNEWGSRPKKLADMYGGGDD 829
Db 781 DYEGSGDAASLSLTSASDQDDYDYLNEWGSRPKKLADMYGGGDD 829

RESULT 14
AD020015
ID ADQ20015 standard; protein; 829 AA.
XX AC ADQ20015;
XX DT 26-AUG-2004 (first entry)
XX DE Human soft tissue sarcoma-upregulated protein - SEQ ID 2835.
XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX OS Homo sapiens.
XX PN WO2004048938-A2.
XX PD 10-JUN-2004.
XX PF 26-NOV-2003; 2003WO-US038193.
XX PR 26-NOV-2002; 2002US-0429739P.
XX (PROT-) PROTEIN DESIGN LABS INC.
XX PA Aziz N, Ginsburg WM, Zlotnik A;
XX PI WPI; 2004-441208/41.
XX DR Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX Example 2; SEQ ID NO 2835; 210pp; English.
XX PS The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX SQ Sequence 829 AA;
Query Match 100.0%; Score 4369; DB 8; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLPRGLASLLLVLCWQAASEPCRAVFRAEVTLGAGAEQPGQALGVFMGCPG 60
Db 1 MGLPRGLASLLLVLCWQAASEPCRAVFRAEVTLGAGAEQPGQALGVFMGCPG 60
QY 61 QEPALFSTDNDDFTVRNGETVQERRSIKERNPLKIPPSKILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDDFTVRNGETVQERRSIKERNPLKIPPSKILRRHKRDWVAPISVPENG 120
QY 121 KGPPFQRLNQLKSNKDRDTKIFYISITGPGADSPPEGVFAVEKETGWLKLNKPLDREBEIAK 180
```

```
Db 121 KGPPFQRLNQLKSNKDRDTKIFYISITGPGADSPPEGVFAVEKETGWLKLNKPLDREBEIAK 180
QY 181 YELFGHAYSENGASVEDPMNISIIVTDQNDHKPKFTQDTFRGSLVLEGLVLPGTSMQVTTAT 240
Db 181 YELFGHAYSENGASVEDPMNISIIVTDQNDHKPKFTQDTFRGSLVLEGLVLPGTSMQVTTAT 240
QY 241 DEDDAIYTYNGVVAYSIHSQEPKDPHDLMTTHRSTGTISVISSSGLDREKVPYETLTIOA 300
Db 241 DEDDAIYTYNGVVAYSIHSQEPKDPHDLMTTHRSTGTISVISSSGLDREKVPYETLTIOA 300
QY 301 TDMGDSGTTTAVAVVEILDANDNAPMFPDQYEAHVPENAVCHVEQRLTVTDLDPNSP 360
Db 301 TDMGDSGTTTAVAVVEILDANDNAPMFPDQYEAHVPENAVCHVEQRLTVTDLDPNSP 360
QY 361 AWRAIYLMGGDGDHFTTTHPESNOGILLTRKGLDFAKQHQHTLYVETNEAPFVLKL 420
Db 361 AWRAIYLMGGDGDHFTTTHPESNOGILLTRKGLDFAKQHQHTLYVETNEAPFVLKL 420
QY 421 PTSTATIVVHVEDVNEAPVFPSPKVVEQEGIPGEPVCVYTAEDPDKENOKISYRIILR 480
Db 421 PTSTATIVVHVEDVNEAPVFPSPKVVEQEGIPGEPVCVYTAEDPDKENOKISYRIILR 480
QY 481 DPAGWLAMPDPSGQVTAVGTLDRDEQFVRNNIYEVWVLMONGSPPTTGTGLTLTLID 540
Db 481 DPAGWLAMPDPSGQVTAVGTLDRDEQFVRNNIYEVWVLMONGSPPTTGTGLTLTLID 540
QY 541 VNDHGPVPRPROITTCNQSPVRHVLNITDKLSPHTSFPAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPRPROITTCNQSPVRHVLNITDKLSPHTSFPAQLTDDSDIYWTAEVNEEGD 600
QY 601 TVVLSLKFKELKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKGFFILPVL 660
Db 601 TVVLSLKFKELKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKGFFILPVL 660
QY 661 GAVLALLFLLLVLLVLRKKRIKEPLLPEDDTRDNVFIYEGEGGEDDYDITQLHR 720
Db 661 GAVLALLFLLLVLLVLRKKRIKEPLLPEDDTRDNVFIYEGEGGEDDYDITQLHR 720
QY 721 GLEARPEVLRNDVAPTIPTMYRPRRPNPDEIGNFIENLKAANTDPTAPPYDILLVF 780
Db 721 GLEARPEVLRNDVAPTIPTMYRPRRPNPDEIGNFIENLKAANTDPTAPPYDILLVF 780
QY 781 DYEGSGDAASLSLTSASDQDDYDYLNEWGSRPKKLADMYGGGDD 829
Db 781 DYEGSGDAASLSLTSASDQDDYDYLNEWGSRPKKLADMYGGGDD 829

RESULT 15
ADP26905
ID ADP26905 standard; protein; 829 AA.
XX AC ADP26905;
XX DT 26-AUG-2004 (first entry)
XX DE Human P-cadherin protein SEQ ID NO:6.
XX KW hair growth modulator; P-cadherin modulator; endocrine; depilatory;
KW gene therapy; antisense therapy; hair growth; alopecia; baldness;
KW unwanted hair growth; hirsutism;
KW hypotrichosis associated with juvenile macular dystrophy; HJMD; human;
KW P-cadherin.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Misc-difference 563 /note= "encoded by CAG"
XX PN EP1428893-A2.
XX PD 16-JUN-2004.
```

XX 10-OCT-2003; 2003EP-00256411.
XX PF
XX PR 15-OCT-2002; 2002US-0418163P.
XX PA (SPRE/) SPRECHER E.
XX PA (BERG/) BERGMAN R.
XX PI Sprecher E, Bergman R;
XX DR WPI; 2004-469945/45.
XX DR N-PSDB; ADP26910.
XX PT Identifying a hair growth modulator for treating alopecia and unwanted
PT hair growth such as hirsutism, comprises identifying a P-cadherin
PT modulator and testing whether the P-cadherin modulator is functional as a
PT hair growth modulator.
XX PS Disclosure; SEQ ID NO 6; 121bp; English.
XX PS The present invention describes a method (M1) for identifying a hair
CC growth modulator. (M1) comprises identifying a P-cadherin modulator, and
CC testing whether the P-cadherin modulator is functional as a hair growth
CC modulator. Also described: (1) a hair growth modulator (I) identified by
CC (M1); and (2) a composition (II) for modulating hair growth, comprising,
CC as an active ingredient, a P-cadherin modulator functional as a hair
CC growth modulator. (I) and (II) have endocrine and depilatory activities,
CC and can be used as hair growth modulators, P-cadherin function
CC modulators, and in gene and antisense therapy. (M1) is useful for
CC identifying a hair growth modulator. (I) is useful in a method of medical
CC treatment. (I) or (II) is useful for modulating hair growth for non-
CC therapeutic cosmetic purposes which involves administering to a subject,
CC (I) or (II). (I) can be used in the manufacture of a medicament for the
CC therapeutic modulation of hair growth. (I) or (II) is useful for treating
CC alopecia (baldness) and unwanted hair growth such as hirsutism. (I) or
CC (II) comprising P-cadherin inducer is useful for correction of hair loss
CC in congenital hypotrichosis associated with juvenile macular dystrophy
CC (HJMD) and other alopecia patients. The present sequence represents human
CC P-cadherin, which is used in the exemplification of the present
XX invention.
XX SQ Sequence 829 AA;

Query Match 100.0%; Score 4369; DB 8; Length 829;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGLPRGLASLLLLQVCWLOCAASEPCRAVPREAEVTLGAGAEQEPQALGKVFMCPCG 60
Db 1 MGLPRGLASLLLLQVCWLOCAASEPCRAVPREAEVTLGAGAEQEPQALGKVFMCPCG 60
Qy 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120
Qy 121 KGPPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREETAK 180
Db 121 KGPPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREETAK 180
Qy 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTPRGSVLEGLPGTSVMQVTAT 240
Db 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTPRGSVLEGLPGTSVMQVTAT 240
Qy 241 DEDDAIYTYNGWAVSIHQEPKDPHDLMFTHIRSTGTISVISSGLDREKYPEYTLTIQA 300
Db 241 DEDDAIYTYNGWAVSIHQEPKDPHDLMFTHIRSTGTISVISSGLDREKYPEYTLTIQA 300
Qy 301 TMDGDGSGTTTAVAVEILDANDNAPMFDPOKYEAHVPENAVGVHVEQRLTVDLDAPNSP 360
Db 301 TMDGDGSGTTTAVAVEILDANDNAPMFDPOKYEAHVPENAVGVHVEQRLTVDLDAPNSP 360
Qy 361 AWRATYLMGGDGDGHFTTTHPESNQGILTRKGLDFEAKNQHTLYVEVTNEAPFVLKL 420
Db 361 AWRATYLMGGDGDGHFTTTHPESNQGILTRKGLDFEAKNQHTLYVEVTNEAPFVLKL 420

Search completed: December 12, 2005, 09:11:16
Job time : 194 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2005, 09:08:03 ; Search time 167 Seconds
(without alignments)
2074.135 Million cell updates/sec

Title: US-10-696-639-39
Perfect score: 4369
Sequence: 1 MGLPRGPLASLLQLQVCWLQ.....NEWGRFKKLADMYGGEDD 829

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4369	100.0	829	3 US-09-905-983-50	Sequence 50, Appl
2	4369	100.0	829	3 US-09-916-849A-1	Sequence 1, Appl
3	4369	100.0	829	3 US-09-975-723A-1	Sequence 1, Appl
4	4369	100.0	829	4 US-10-174-677-21	Sequence 21, Appl
5	4369	100.0	829	4 US-10-158-123-1	Sequence 1, Appl
6	4369	100.0	829	4 US-10-295-027-783	Sequence 783, Appl
7	4369	100.0	829	4 US-10-295-027-800	Sequence 800, Appl
8	4369	100.0	829	4 US-10-295-027-863	Sequence 863, Appl
9	4369	100.0	829	4 US-10-295-027-896	Sequence 896, Appl
10	4369	100.0	829	4 US-10-058-270A-126	Sequence 126, Appl
11	4369	100.0	829	5 US-10-643-795A-145	Sequence 145, Appl
12	4369	100.0	829	5 US-10-578-160A-6	Sequence 6, Appl
13	4369	100.0	829	5 US-10-723-860-2835	Sequence 2835, Ap
14	4369	100.0	829	5 US-10-696-639-39	Sequence 39, Appl
15	4369	100.0	829	5 US-10-948-518-145	Sequence 145, Appl
16	4369	100.0	829	5 US-10-756-149-5498	Sequence 5498, Ap
17	4369	100.0	829	6 US-11-037-713-22	Sequence 22, Appl
18	4361	99.8	829	4 US-10-301-822-18	Sequence 18, Appl
19	4361	99.8	829	4 US-10-295-027-338	Sequence 338, Appl
20	4361	99.8	829	4 US-10-295-027-1246	Sequence 1246, Ap
21	4361	99.8	829	4 US-10-087-080-14	Sequence 14, Appl
22	4361	99.8	829	4 US-10-229-345-16	Sequence 16, Appl
23	4361	99.8	829	4 US-10-274-177-16	Sequence 16, Appl
24	4361	99.8	829	4 US-10-650-112-16	Sequence 16, Appl
25	4361	99.8	829	4 US-10-712-124-86	Sequence 86, Appl
26	4361	99.8	829	6 US-11-041-788-16	Sequence 16, Appl
27	2420	55.4	882	4 US-10-097-340-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1
US-09-905-983-50
; Sequence 50, Application US/09905983
; Patent No. US20020045591A1
; GENERAL INFORMATION:
; APPLICANT: Geiger, Benjamin
; APPLICANT: Ben-Ze'ev, Avri
; APPLICANT: Sadot, Eilat
; TITLE OF INVENTION: METHODS AND THERAPEUTIC COMPOSITIONS FOR TREATING CANCER
; FILE REFERENCE: 01/22326
; CURRENT APPLICATION NUMBER: US/09/905,983
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 50
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-983-50

Query Match	100.0%;	Score 4369;	DB 3;	Length 829;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 829;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MGLPRGPLASLLQLQVCWLQCAASPCRAVFREAEVTL	EAGGAQEPGQALGKVFNGCPG	60
Db	1	MGLPRGPLASLLQLQVCWLQCAASPCRAVFREAEVTL	EAGGAQEPGQALGKVFNGCPG	60
QY	61	QEPALFSTDNDDFTVRNGETVQERSLKERNPLKIPPSKRI	LRHKRWVAVPISVPENG	120
Db	61	QEPALFSTDNDDFTVRNGETVQERSLKERNPLKIPPSKRI	LRHKRWVAVPISVPENG	120
QY	121	KGPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETG	WLLLNKPLDREIAK	180
Db	121	KGPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETG	WLLLNKPLDREIAK	180
QY	181	YELFGHVAHSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGS	VLGVLPGTSMQVTTAT	240
Db	181	YELFGHVAHSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGS	VLGVLPGTSMQVTTAT	240
QY	241	DDDDAIYTVNGVAVSIHQEPKDPHDLMTFTHRSITSVIS	SSGLDREKVPYETITOA	300
Db	241	DDDDAIYTVNGVAVSIHQEPKDPHDLMTFTHRSITSVIS	SSGLDREKVPYETITOA	300
QY	301	TDMDGSGSTTTTAVAVVEILLDANDNAPMFPQKYEAHPENAVG	HEVQRITVTDLDAPNSP	360
Db	301	TDMDGSGSTTTTAVAVVEILLDANDNAPMFPQKYEAHPENAVG	HEVQRITVTDLDAPNSP	360
QY	361	AWRAYLYIMGDDGDHFTTTHPESNQGLLTRKGLDFAKQHTLY	VEVTHNEAPVFLKL	420

```
Db 361 AWRATYLMGGDDGDHFTITTHPESNOGILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
Qy 421 PTSTATIIVHVEDVNEAPFVFPSPKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480
Db 421 PTSTATIIVHVEDVNEAPFVFPSPKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480
Qy 481 DPAGWLAMDPSQGVTAAGTGLDREDEQFVRNNIYEVNVLAMDNPSPTTGTGTLTLLTLD 540
Db 481 DPAGWLAMDPSQGVTAAGTGLDREDEQFVRNNIYEVNVLAMDNPSPTTGTGTLTLLTLD 540
Qy 541 VNDHGPVPEPRQITICNOSPVRHVLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPRQITICNOSPVRHVLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSSKKFLKQDHYDHLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Db 601 TVVLSSKKFLKQDHYDHLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Qy 661 GAVLALLFLLVLLVLLVLRKKRKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 720
Db 661 GAVLALLFLLVLLVLLVLRKKRKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 720
Qy 721 GLEARPEVLRNDVAPTIIPTPMYRPRPANDPDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVLRNDVAPTIIPTPMYRPRPANDPDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
Qy 781 DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGDD 829
Db 781 DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGDD 829

RESULT 2
US-09-916-849A-1
; Sequence 1, Application US/09916849A
; Publication No. US20030086934A1
; GENERAL INFORMATION:
; APPLICANT: Bostein, et al.
; TITLE OF INVENTION: Basal Markers in Breast Cancer and Related Reagents
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 2002850-0024
; CURRENT APPLICATION NUMBER: US/09/916.849A
; CURRENT FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Sequence of
; OTHER INFORMATION: Cadherin 3
US-09-916-849A-1

Query Match 100.0%; Score 4369; DB 3; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGPLASLLILQVCWLQCAASEPCRAVFREAEVTLAAGAEQEPQALGKVFMCPCG 60
Db 1 MGLPRGPLASLLILQVCWLQCAASEPCRAVFREAEVTLAAGAEQEPQALGKVFMCPCG 60
Qy 61 QEPALFSTNDNDFTRVNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTNDNDFTRVNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120
Qy 121 KGPPFPQRLNQLKSNKDRDTKIFYGITGPGADSPPEGVFAVEKETGWLILNKLPLDREBIAK 180
Db 121 KGPPFPQRLNQLKSNKDRDTKIFYGITGPGADSPPEGVFAVEKETGWLILNKLPLDREBIAK 180
Qy 181 YELFGHAVSENGASVEDPMNIIIVTDNDHKPKFTQDTPFRGSVLEGLVPGTSMQVTAT 240
Db 181 YELFGHAVSENGASVEDPMNIIIVTDNDHKPKFTQDTPFRGSVLEGLVPGTSMQVTAT 240
```

```
Qy 241 DEDDAIYTYNGVVAYSIHQSEPKDPHDLMTFTHRSTGTISVISSSGLDREKVPYTLTIOA 300
Db 241 DEDDAIYTYNGVVAYSIHQSEPKDPHDLMTFTHRSTGTISVISSSGLDREKVPYTLTIOA 300
Qy 301 TMDGSGSTTTAVAVVEIILDANDNAPMFDPQKYEAHVPENAVGEVQRLTVTDLDPNSP 360
Db 301 TMDGSGSTTTAVAVVEIILDANDNAPMFDPQKYEAHVPENAVGEVQRLTVTDLDPNSP 360
Qy 361 AWRATYLMGGDDGDHFTITTHPESNOGILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
Db 361 AWRATYLMGGDDGDHFTITTHPESNOGILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
Qy 421 PTSTATIIVHVEDVNEAPFVFPSPKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480
Db 421 PTSTATIIVHVEDVNEAPFVFPSPKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480
Qy 481 DPAGWLAMDPSQGVTAAGTGLDREDEQFVRNNIYEVNVLAMDNPSPTTGTGTLTLLTLD 540
Db 481 DPAGWLAMDPSQGVTAAGTGLDREDEQFVRNNIYEVNVLAMDNPSPTTGTGTLTLLTLD 540
Qy 541 VNDHGPVPEPRQITICNOSPVRHVLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPRQITICNOSPVRHVLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSSKKFLKQDHYDHLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Db 601 TVVLSSKKFLKQDHYDHLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Qy 661 GAVLALLFLLVLLVLLVLRKKRKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 720
Db 661 GAVLALLFLLVLLVLLVLRKKRKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 720
Qy 721 GLEARPEVLRNDVAPTIIPTPMYRPRPANDPDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVLRNDVAPTIIPTPMYRPRPANDPDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
Qy 781 DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGDD 829
Db 781 DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGDD 829

RESULT 3
US-09-975-723A-1
; Sequence 1, Application US/09975723A
; Publication No. US20030108529A1
; GENERAL INFORMATION:
; APPLICANT: Nackman, Gary
; APPLICANT: Foty, Ramsey
; TITLE OF INVENTION: Improvement of Endothelial Cell-Cell
; TITLE OF INVENTION: Cohesion
; FILE REFERENCE: 601-1-101N
; CURRENT APPLICATION NUMBER: US/09/975.723A
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/241,216
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/243,693
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-975-723A-1

Query Match 100.0%; Score 4369; DB 3; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGPLASLLILQVCWLQCAASEPCRAVFREAEVTLAAGAEQEPQALGKVFMCPCG 60
Db 1 MGLPRGPLASLLILQVCWLQCAASEPCRAVFREAEVTLAAGAEQEPQALGKVFMCPCG 60
```

Qy	61	QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDWVWVAPISVPENG	120
Db	61	QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDWVWVAPISVPENG	120
Qy	121	KGPPQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREEIAK	180
Db	121	KGPPQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREEIAK	180
Qy	181	YELFGHAVSNGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSLVGLVPGTSMQVTTAT	240
Db	181	YELFGHAVSNGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSLVGLVPGTSMQVTTAT	240
Qy	241	DEDDAIYTYNGVAVSIHSQEPKDPHDLMTIHRSTGTISVISGGLDREKVPYTYLTIQA	300
Db	241	DEDDAIYTYNGVAVSIHSQEPKDPHDLMTIHRSTGTISVISGGLDREKVPYTYLTIQA	300
Qy	301	TMMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAVHPENAVGHEVQRLTVTDLDPNSP	360
Db	301	TMMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAVHPENAVGHEVQRLTVTDLDPNSP	360
Qy	361	AWRATYLMGGDDGDHFTITTHPESNQILITTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420
Db	361	AWRATYLMGGDDGDHFTITTHPESNQILITTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420
Qy	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEVQEGIPTGEPVCVYTAEDPDKENQKISYRIILR	480
Db	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEVQEGIPTGEPVCVYTAEDPDKENQKISYRIILR	480
Qy	481	DPAGWLAMPDPSGOVTAVGTLDRDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLLTLD	540
Db	481	DPAGWLAMPDPSGOVTAVGTLDRDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLLTLD	540
Qy	541	VNDHGPVPEPRQITICNOSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD	600
Db	541	VNDHGPVPEPRQITICNOSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD	600
Qy	601	TVVLSLKKFLKQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILPVL	660
Db	601	TVVLSLKKFLKQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILPVL	660
Qy	661	GAVLALLFLLVLLVLLVLRKKRKEPLLLPEDDTRDNVFFYEGEGGEDDQDYDTQLHR	720
Db	661	GAVLALLFLLVLLVLLVLRKKRKEPLLLPEDDTRDNVFFYEGEGGEDDQDYDTQLHR	720
Qy	721	GLEAPRWLRNDVAPITIIPTMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF	780
Db	721	GLEAPRWLRNDVAPITIIPTMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF	780
Qy	781	DYEGSGDAASLSLSTSSASDQDQDYDYLNEWGSRFKKLADMYGGGEDD	829
Db	781	DYEGSGDAASLSLSTSSASDQDQDYDYLNEWGSRFKKLADMYGGGEDD	829

RESULT 4
US-10-174-677-21
; Sequence 21, Application US/10174677
; Publication No. US20030190704A1
; GENERAL INFORMATION:
; APPLICANT: Xie, Ting
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ANCHORING STEM CELLS IN A MICROENVIR
; FILE REFERENCES: 40716(IP-012)
; CURRENT APPLICATION NUMBER: US/10/174,677
; CURRENT FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 21
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-677-21

Query Match 100.0%; Score 4369; DB 4; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;

		Matches	829;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MGLPRGPLASLLLLLQVCMWLQCAASPPCRAVFPREAEVTLLEAGGABQEPGQALGKVFMGCPG	60								
Db	1	MGLPRGPLASLLLLLQVCMWLQCAASPPCRAVFPREAEVTLLEAGGABQEPGQALGKVFMGCPG	60								
Qy	61	QBPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDWVWVAPISVPENG	120								
Db	61	QBPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDWVWVAPISVPENG	120								
Qy	121	KGPPQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREEIAK	180								
Db	121	KGPPQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREEIAK	180								
Qy	181	YELFGHAVSNGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSLVGLVPGTSMQVTTAT	240								
Db	181	YELFGHAVSNGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSLVGLVPGTSMQVTTAT	240								
Qy	241	DEDDAIYTYNGVAVSIHSQEPKDPHDLMTIHRSTGTISVISGGLDREKVPYTYLTIQA	300								
Db	241	DEDDAIYTYNGVAVSIHSQEPKDPHDLMTIHRSTGTISVISGGLDREKVPYTYLTIQA	300								
Qy	301	TMMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAVHPENAVGHEVQRLTVTDLDPNSP	360								
Db	301	TMMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAVHPENAVGHEVQRLTVTDLDPNSP	360								
Qy	361	AWRATYLMGGDDGDHFTITTHPESNQILITTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420								
Db	361	AWRATYLMGGDDGDHFTITTHPESNQILITTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420								
Qy	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEVQEGIPTGEPVCVYTAEDPDKENQKISYRIILR	480								
Db	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEVQEGIPTGEPVCVYTAEDPDKENQKISYRIILR	480								
Qy	481	DPAGWLAMPDPSGOVTAVGTLDRDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLLTLD	540								
Db	481	DPAGWLAMPDPSGOVTAVGTLDRDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLLTLD	540								
Qy	541	VNDHGPVPEPRQITICNOSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD	600								
Db	541	VNDHGPVPEPRQITICNOSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD	600								
Qy	601	TVVLSLKKFLKQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILPVL	660								
Db	601	TVVLSLKKFLKQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILPVL	660								
Qy	661	GAVLALLFLLVLLVLLVLRKKRKEPLLLPEDDTRDNVFFYEGEGGEDDQDYDTQLHR	720								
Db	661	GAVLALLFLLVLLVLLVLRKKRKEPLLLPEDDTRDNVFFYEGEGGEDDQDYDTQLHR	720								
Qy	721	GLEAPRWLRNDVAPITIIPTMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF	780								
Db	721	GLEAPRWLRNDVAPITIIPTMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF	780								
Qy	781	DYEGSGDAASLSLSTSSASDQDQDYDYLNEWGSRFKKLADMYGGGEDD	829								
Db	781	DYEGSGDAASLSLSTSSASDQDQDYDYLNEWGSRFKKLADMYGGGEDD	829								

RESULT 5
US-10-158-123-1
; Sequence 1, Application US/10158123
; Publication No. US20030194406A1
; GENERAL INFORMATION:
; APPLICANT: Reinhard, Christoph
; APPLICANT: Klingner, Julie
; APPLICANT: Jefferson, Ann
; APPLICANT: Escobedo, Jaime
; APPLICANT: Randazzo, Fillipo
; APPLICANT: Winter, Jill
; APPLICANT: Goodson, Robert
; APPLICANT: Qi, Weimin
; TITLE OF INVENTION: p-Cadherin as a Target for Anti-Cancer


```
; TITLE OF INVENTION: Therapy
; FILE REFERENCE: 35784/258994
; CURRENT APPLICATION NUMBER: US/10/158,123
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/294,225
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-158-123-1

Query Match      100.0%; Score 4369; DB 4; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLEAGGAEQEPQALGKVFMGCPG 60
Db 1 MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLEAGGAEQEPQALGKVFMGCPG 60

Qy 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120

Qy 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120

Qy 121 KGPFPQRLNQLKSKNDRDKIFYSITGPGADSPPEGVFAVEKETGWLKLLNKLPLDREIAK 180
Db 121 KGPFPQRLNQLKSKNDRDKIFYSITGPGADSPPEGVFAVEKETGWLKLLNKLPLDREIAK 180

Qy 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTPFGSVLEGVLPGTSMQVTAT 240
Db 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTPFGSVLEGVLPGTSMQVTAT 240

Qy 241 DEDDAIYTYNGVAVSIHSQEPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
Db 241 DEDDAIYTYNGVAVSIHSQEPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300

; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 783
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-295-027-783

Query Match      100.0%; Score 4369; DB 4; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLEAGGAEQEPQALGKVFMGCPG 60
Db 1 MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLEAGGAEQEPQALGKVFMGCPG 60

Qy 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120

Qy 121 KGPFPQRLNQLKSKNDRDKIFYSITGPGADSPPEGVFAVEKETGWLKLLNKLPLDREIAK 180
Db 121 KGPFPQRLNQLKSKNDRDKIFYSITGPGADSPPEGVFAVEKETGWLKLLNKLPLDREIAK 180

Qy 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTPFGSVLEGVLPGTSMQVTAT 240
Db 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTPFGSVLEGVLPGTSMQVTAT 240

Qy 241 DEDDAIYTYNGVAVSIHSQEPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
Db 241 DEDDAIYTYNGVAVSIHSQEPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
```

Db 241 DEDDAIYVNGVAYSHSQEPKDPHDLMTIHRSTGTISVSSGLDRKVPYTLTIOA 300
Qy 301 TMDGSGSTTTAVAVVEILDANDNAPMDFPQKYEAHVPENAVGHVQRLTVDLDAPNSP 360
Db 301 TMDGSGSTTTAVAVVEILDANDNAPMDFPQKYEAHVPENAVGHVQRLTVDLDAPNSP 360
Qy 361 AWRATYLMGDDGDHFTITTHPESNQILTKDLSPHSPFOAQLTDDSDIYWTAEVNEEGL 420
Db 361 AWRATYLMGDDGDHFTITTHPESNQILTKDLSPHSPFOAQLTDDSDIYWTAEVNEEGL 420
Qy 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTEGEPVCVYTAEDPDKENOKISYRIILR 480
Db 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTEGEPVCVYTAEDPDKENOKISYRIILR 480
Qy 481 DPAGWLAMPDSDGQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGTLTLLTID 540
Db 481 DPAGWLAMPDSDGQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGTLTLLTID 540
Qy 541 VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPHSPFOAQLTDDSDIYWTAEVNEEGL 600
Db 541 VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPHSPFOAQLTDDSDIYWTAEVNEEGL 600
Qy 601 TVVLSLKFKLQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFLLPVL 660
Db 601 TVVLSLKFKLQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFLLPVL 660
Qy 661 GAVLALLFLLVLLVLLVLRKKRIKEPLLLPEDDTRDNVFFYEGEGEEDQDYDITQLHR 720
Db 661 GAVLALLFLLVLLVLLVLRKKRIKEPLLLPEDDTRDNVFFYEGEGEEDQDYDITQLHR 720
Qy 721 GLEAREVVLNRNDVAPTIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDILLVF 780
Db 721 GLEAREVVLNRNDVAPTIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDILLVF 780
Qy 781 DYEGSGSDAASLSLTSASDQDQDYDYLNEGWSRKKLADMYGGGEDD 829
Db 781 DYEGSGSDAASLSLTSASDQDQDYDYLNEGWSRKKLADMYGGGEDD 829

RESULT 7

US-10-295-027-800
; Sequence 800, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Aziz, Natasha
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08

; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 800
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-295-027-800

Query Match 100.0%; Score 4369; DB 4; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGPLASLLLLQVCMQLCAASBPCRAVFPREAEVTLLEAGGAQEPGQALGKVFMGCPG 60
Db 1 MGLPRGPLASLLLLQVCMQLCAASBPCRAVFPREAEVTLLEAGGAQEPGQALGKVFMGCPG 60
Qy 61 QEPALFSTDNDDFTVRNGETVQERSLKERNPLKIPPSKRIILRRHKRDVWVAPISVPENG 120
Db 61 QEPALFSTDNDDFTVRNGETVQERSLKERNPLKIPPSKRIILRRHKRDVWVAPISVPENG 120
Qy 121 KGPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREEIAK 180
Db 121 KGPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREEIAK 180
Qy 181 YELFHGAVSENGASVEDPMNISIIVTQNDHDKPKTQDTFRGSLVLEGLPGTSVMQVTAT 240
Db 181 YELFHGAVSENGASVEDPMNISIIVTQNDHDKPKTQDTFRGSLVLEGLPGTSVMQVTAT 240
Qy 241 DEDDAIYVNGVAYSHSQEPKDPHDLMTIHRSTGTISVSSGLDRKVPYTLTIOA 300
Db 241 DEDDAIYVNGVAYSHSQEPKDPHDLMTIHRSTGTISVSSGLDRKVPYTLTIOA 300
Qy 301 TMDGSGSTTTAVAVVEILDANDNAPMDFPQKYEAHVPENAVGHVQRLTVDLDAPNSP 360
Db 301 TMDGSGSTTTAVAVVEILDANDNAPMDFPQKYEAHVPENAVGHVQRLTVDLDAPNSP 360
Qy 361 AWRATYLMGDDGDHFTITTHPESNQILTKDLSPHSPFOAQLTDDSDIYWTAEVNEEGL 420
Db 361 AWRATYLMGDDGDHFTITTHPESNQILTKDLSPHSPFOAQLTDDSDIYWTAEVNEEGL 420
Qy 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTEGEPVCVYTAEDPDKENOKISYRIILR 480
Db 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTEGEPVCVYTAEDPDKENOKISYRIILR 480
Qy 481 DPAGWLAMPDSDGQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGTLTLLTID 540
Db 481 DPAGWLAMPDSDGQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGTLTLLTID 540
Qy 541 VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPHSPFOAQLTDDSDIYWTAEVNEEGL 600
Db 541 VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPHSPFOAQLTDDSDIYWTAEVNEEGL 600
Qy 601 TVVLSLKFKLQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFLLPVL 660
Db 601 TVVLSLKFKLQDQTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFLLPVL 660
Qy 661 GAVLALLFLLVLLVLLVLRKKRIKEPLLLPEDDTRDNVFFYEGEGEEDQDYDITQLHR 720
Db 661 GAVLALLFLLVLLVLLVLRKKRIKEPLLLPEDDTRDNVFFYEGEGEEDQDYDITQLHR 720
Qy 721 GLEAREVVLNRNDVAPTIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDILLVF 780
Db 721 GLEAREVVLNRNDVAPTIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDILLVF 780
Qy 781 DYEGSGSDAASLSLTSASDQDQDYDYLNEGWSRKKLADMYGGGEDD 829
Db 781 DYEGSGSDAASLSLTSASDQDQDYDYLNEGWSRKKLADMYGGGEDD 829

```
Db      781 DYEGSGSAAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGDD 829

RESULT 8
US-10-295-027-863
; Sequence 863, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 863
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-863

Query Match      100.0%; Score 4369; DB 4; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MCLPRGPLASLLLOVCWLQCAASPCRAVREABVTLEAGAEQPCQALGKVFMC 60
Db      1 MCLPRGPLASLLLOVCWLQCAASPCRAVREABVTLEAGAEQPCQALGKVFMC 60

Qy      61 QEPALFSTDNDDFTVRNGETVQERSLKERPLKIFPSKRILRRHKRDWVVAISVPENG 120
Db      61 QEPALFSTDNDDFTVRNGETVQERSLKERPLKIFPSKRILRRHKRDWVVAISVPENG 120

Qy      121 KGPPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVKECTGWLNLNKLPLDREETAK 180
Db      121 KGPPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVKECTGWLNLNKLPLDREETAK 180

Qy      181 YELFCHAVSENGASVEDPMNLSIIYTDQNDHKPKFTQDTFRGSVLGVLPGTSMQVNTAT 240
Db      181 YELFCHAVSENGASVEDPMNLSIIYTDQNDHKPKFTQDTFRGSVLGVLPGTSMQVNTAT 240

Qy      241 DEDDAIYTYNGVAYSIHSQEPKDPHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300

Db      241 DEDDAIYTYNGVAYSIHSQEPKDPHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300

301    TDMGDSGTTTAVAVVEILDANDNAPMEDPQKYEAHVPENAVGHEVQRLTVTDLDAENSP 360
301    TDMGDSGTTTAVAVVEILDANDNAPMEDPQKYEAHVPENAVGHEVQRLTVTDLDAENSP 360
361    AWRATYLMGGDDGDHFTITTHPESNQGLITTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
361    AWRATYLMGGDDGDHFTITTHPESNQGLITTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
421    PTSTATIVHVEDVNEAPFVPPSKVVEVOEGIPTEGPEVCVYTAEDPDKENQKISYRILR 480
421    PTSTATIVHVEDVNEAPFVPPSKVVEVOEGIPTEGPEVCVYTAEDPDKENQKISYRILR 480
481    DRAGWLAMPDSCGVTAAGTGLDREDEQFVRNNIYEVVWLAMDNGSPPTTGTGTLTLLID 540
481    DRAGWLAMPDSCGVTAAGTGLDREDEQFVRNNIYEVVWLAMDNGSPPTTGTGTLTLLID 540
541    VNDHGPVPEPROITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
541    VNDHGPVPEPROITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
601    TVVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILPVL 660
601    TVVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILPVL 660
661    GAVLALLFLLLVLLVLRKKRKEQLLPEDDTRDNVFFYEGEGEEDQDYYDTOLHR 720
661    GAVLALLFLLLVLLVLRKKRKEQLLPEDDTRDNVFFYEGEGEEDQDYYDTOLHR 720
721    GLEARPEVLRNDVAPTIIPTPMYRPRPANDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
721    GLEARPEVLRNDVAPTIIPTPMYRPRPANDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
781    DYEGSGSAAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGDD 829
781    DYEGSGSAAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGDD 829

RESULT 9
US-10-295-027-896
; Sequence 896, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
```

;; PRIOR FILING DATE: 2002-01-08
;; PRIOR APPLICATION NUMBER: US 60/347,349
;; PRIOR FILING DATE: 2002-01-10
;; PRIOR APPLICATION NUMBER: US 60/355,250
;; PRIOR FILING DATE: 2002-02-08
;; PRIOR APPLICATION NUMBER: US 60/356,714
;; PRIOR FILING DATE: 2002-02-13
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 1386
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 896
;; LENGTH: 829
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-295-027-896

Query Match 100.0%; Score 4369; DB 4; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLPRGLASLLQLQVCWLCQAASEPCRAVFREAEVTLLEAGGAEQPGQALGKVFPGCG 60
DB 1 MGLPRGLASLLQLQVCWLCQAASEPCRAVFREAEVTLLEAGGAEQPGQALGKVFPGCG 60
QY 61 QEPALFSTNDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG 120
DB 61 QEPALFSTNDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG 120
QY 121 KGPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLLNKLPLDREEIAK 180
DB 121 KGPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLLNKLPLDREEIAK 180
QY 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLGVLPGTSVMQVAT 240
DB 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLGVLPGTSVMQVAT 240
QY 241 DEDDAIYTYNGVAVSIHSGEPKDPDLMTTHRSTGTISVSSGLDREKVPYTLTIOA 300
DB 241 DEDDAIYTYNGVAVSIHSGEPKDPDLMTTHRSTGTISVSSGLDREKVPYTLTIOA 300
QY 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAVHPENAVGHEVQRLTVTDLDPNSP 360
DB 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAVHPENAVGHEVQRLTVTDLDPNSP 360
QY 361 AWRATYLVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENOKISYRIILR 480
DB 361 AWRATYLVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENOKISYRIILR 480
QY 481 DRAGWLAMPDSCGVATGTLDRDQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLTLD 540
DB 481 DRAGWLAMPDSCGVATGTLDRDQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLTLD 540
QY 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
DB 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
QY 601 TVVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKGFFLPVL 660
DB 601 TVVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKGFFLPVL 660
QY 661 GAVLALLFLVLVLLVLRKKRKIKPELLIPEDTRDNVYFYGEGGEBEDQYDIQTLHR 720
DB 661 GAVLALLFLVLVLLVLRKKRKIKPELLIPEDTRDNVYFYGEGGEBEDQYDIQTLHR 720
QY 721 GLEARPEVLRNDVAPTIIPTPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
DB 721 GLEARPEVLRNDVAPTIIPTPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
QY 781 DYEGSGDAASLSLTSASDQDDYDYLNEWGSRFKKLADMYGGGEDD 829

DB 781 DYEGSGDAASLSLTSASDQDDYDYLNEWGSRFKKLADMYGGGEDD 829

RESULT 10
US-10-058-270A-126
;; Sequence 126, Application US/10058270A
;; Publication No. US20040029114A1
;; GENERAL INFORMATION:
;; APPLICANT: Mack, David H.
;; APPLICANT: Gish, Kurt C.
;; APPLICANT: Afar, Daniel
;; APPLICANT: Eos Biotechnology, Inc.
;; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and
;; FILE REFERENCE: 018501-005210US
;; CURRENT APPLICATION NUMBER: US/10/058,270A
;; CURRENT FILING DATE: 2002-01-24
;; PRIOR APPLICATION NUMBER: US 60/263,965
;; PRIOR FILING DATE: 2001-01-24
;; PRIOR APPLICATION NUMBER: US 60/265,928
;; PRIOR FILING DATE: 2001-02-02
;; PRIOR APPLICATION NUMBER: US 09/829,472
;; PRIOR FILING DATE: 2001-04-09
;; PRIOR APPLICATION NUMBER: US 60/282,698
;; PRIOR FILING DATE: 2001-04-09
;; PRIOR APPLICATION NUMBER: US 60/288,590
;; PRIOR FILING DATE: 2001-05-04
;; PRIOR APPLICATION NUMBER: US 60/294,443
;; PRIOR FILING DATE: 2001-05-29
;; NUMBER OF SEQ ID NOS: 141
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 126
;; LENGTH: 829
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-058-270A-126

Query Match 100.0%; Score 4369; DB 4; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLPRGLASLLQLQVCWLCQAASEPCRAVFREAEVTLLEAGGAEQPGQALGKVFPGCG 60
DB 1 MGLPRGLASLLQLQVCWLCQAASEPCRAVFREAEVTLLEAGGAEQPGQALGKVFPGCG 60
QY 61 QEPALFSTNDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG 120
DB 61 QEPALFSTNDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG 120
QY 121 KGPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLLNKLPLDREEIAK 180
DB 121 KGPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLLNKLPLDREEIAK 180
QY 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLGVLPGTSVMQVAT 240
DB 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLGVLPGTSVMQVAT 240
QY 241 DEDDAIYTYNGVAVSIHSGEPKDPDLMTTHRSTGTISVSSGLDREKVPYTLTIOA 300
DB 241 DEDDAIYTYNGVAVSIHSGEPKDPDLMTTHRSTGTISVSSGLDREKVPYTLTIOA 300
QY 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAVHPENAVGHEVQRLTVTDLDPNSP 360
DB 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAVHPENAVGHEVQRLTVTDLDPNSP 360
QY 361 AWRATYLVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENOKISYRIILR 480
DB 361 AWRATYLVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENOKISYRIILR 480
QY 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENOKISYRIILR 480
DB 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENOKISYRIILR 480

Qy	481	DPAGWLAMPDPSGOVTAVGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLLID	540
Db	481	DPAGWLAMPDPSGOVTAVGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLLID	540
Qy	541	VNDHGPVPEPRQITICNQSPVRHVLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD	600
Db	541	VNDHGPVPEPRQITICNQSPVRHVLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD	600
Qy	601	TVVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL	660
Db	601	TVVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL	660
Qy	661	GAVALALLFLLLVLLVLLVLRKKRKKIKEPLLLPDDTRDNVFFYGGEGGEEQDDYDITOLHR	720
Db	661	GAVALALLFLLLVLLVLLVLRKKRKKIKEPLLLPDDTRDNVFFYGGEGGEEQDDYDITOLHR	720
Qy	721	GLEARPEVLRNDVAPTIIPTPMYRPRPANDPEIGNFIIENLKAANTDPTAPPYDTLLVF	780
Db	721	GLEARPEVLRNDVAPTIIPTPMYRPRPANDPEIGNFIIENLKAANTDPTAPPYDTLLVF	780
Qy	781	DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGDD	829
Db	781	DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGDD	829
RESULT 11			
US-10-643-795A-145			
; Sequence 145, Application US/10643795A			
; Publication No. US20040241703A1			
; GENERAL INFORMATION:			
; APPLICANT: FREDERIC J. DESAUVAGE			
; APPLICANT: GRETCHEN PRANTZ			
; APPLICANT: KENNETH J. HILLAN			
; APPLICANT: PAUL POLAKIS			
; APPLICANT: ANDREW POLSON			
; APPLICANT: VICTORIA SMITH			
; APPLICANT: SUSAN D. SPENCER			
; APPLICANT: THOMAS D. WU			
; APPLICANT: ZEMIN ZHANG			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND			
; TITLE OF INVENTION: TREATMENT OF TUMOR			
; FILE REFERENCE: P5026R1-US			
; CURRENT APPLICATION NUMBER: US/10/643,795A			
; CURRENT FILING DATE: 2003-08-19			
; PRIOR FILING DATE: 2002-08-19			
; PRIOR APPLICATION NUMBER: US 60/404,809			
; PRIOR FILING DATE: 2002-08-19			
; PRIOR APPLICATION NUMBER: US 60/405,645			
; PRIOR FILING DATE: 2002-08-21			
; PRIOR APPLICATION NUMBER: US 60/413,192			
; PRIOR FILING DATE: 2002-09-23			
; PRIOR APPLICATION NUMBER: US 60/419,008			
; PRIOR FILING DATE: 2002-10-15			
; PRIOR APPLICATION NUMBER: US 60/426,847			
; PRIOR FILING DATE: 2002-11-15			
; PRIOR APPLICATION NUMBER: US 60/484,959			
; PRIOR FILING DATE: 2003-07-02			
; NUMBER OF SEQ ID NOS: 158			
; SEQ ID NO 145			
; LENGTH: 829			
; TYPE: PRT			
; ORGANISM: Homo sapien			
US-10-643-795A-145			
Query Match 100.0%; Score 4369; DB 5; Length 829;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MGLPRGPLASLLLLQVCHLQCAASEPCRAVFEAEVTLLEAGAGQEPQOALGKVMGCPG	60
Db	1	MGLPRGPLASLLLLQVCHLQCAASEPCRAVFEAEVTLLEAGAGQEPQOALGKVMGCPG	60
Qy	61	QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHRKRDWVAPISVPENG	120

Db	61	QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHRKRDWVAPISVPENG	120
Qy	121	KGFPFORLNQKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDRBEIAK	180
Db	121	KGFPFORLNQKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDRBEIAK	180
Qy	181	YELFGHAVSENGASVEDPMNISIIVTDQNDHDKPKFTQDTPRGSVLGVLPGTSMVQVAT	240
Db	181	YELFGHAVSENGASVEDPMNISIIVTDQNDHDKPKFTQDTPRGSVLGVLPGTSMVQVAT	240
Qy	241	DEDDAIYTYNGVAVYSIHQOEKPDHLMFTIHRSTGTISVISGGLDREKVPYTLTIOA	300
Db	241	DEDDAIYTYNGVAVYSIHQOEKPDHLMFTIHRSTGTISVISGGLDREKVPYTLTIOA	300
Qy	301	TDMDGSGSTTTAVAVVEILDANDANPMDPOKQEAHVPENAVGHVQRLTVTDLDPNSP	360
Db	301	TDMDGSGSTTTAVAVVEILDANDANPMDPOKQEAHVPENAVGHVQRLTVTDLDPNSP	360
Qy	361	AWRATYILMGDDGDHFTITTHPESNQGLITTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420
Db	361	AWRATYILMGDDGDHFTITTHPESNQGLITTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420
Qy	421	PTSTATIVVHVEDVNEAPVFPSPKVVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR	480
Db	421	PTSTATIVVHVEDVNEAPVFPSPKVVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR	480
Qy	481	DPAGWLAMPDPSGOVTAVGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLLID	540
Db	481	DPAGWLAMPDPSGOVTAVGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLLID	540
Qy	541	VNDHGPVPEPRQITICNQSPVRHVLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD	600
Db	541	VNDHGPVPEPRQITICNQSPVRHVLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD	600
Qy	601	TVVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL	660
Db	601	TVVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL	660
Qy	661	GAVALALLFLLLVLLVLLVLRKKRKKIKEPLLLPDDTRDNVFFYGGEGGEEQDDYDITOLHR	720
Db	661	GAVALALLFLLLVLLVLLVLRKKRKKIKEPLLLPDDTRDNVFFYGGEGGEEQDDYDITOLHR	720
Qy	721	GLEARPEVLRNDVAPTIIPTPMYRPRPANDPEIGNFIIENLKAANTDPTAPPYDTLLVF	780
Db	721	GLEARPEVLRNDVAPTIIPTPMYRPRPANDPEIGNFIIENLKAANTDPTAPPYDTLLVF	780
Qy	781	DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGDD	829
Db	781	DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGDD	829
RESULT 12			
US-10-678-160A-6			
; Sequence 6, Application US/10678160A			
; Publication No. US20040247555A1			
; GENERAL INFORMATION:			
; APPLICANT: Sprecher, Eli			
; APPLICANT: Bergman, Reuven			
; TITLE OF INVENTION: METHODS OF AND COMPOSITIONS FOR MODULATING HAIR GROWTH VIA			
; TITLE OF INVENTION: P-CADHERIN MODULATORS			
; FILE REFERENCE: 26465			
; CURRENT APPLICATION NUMBER: US/10/678,160A			
; CURRENT FILING DATE: 2003-10-06			
; NUMBER OF SEQ ID NOS: 75			
; SOFTWARE: Patent in version 3.1			
; SEQ ID NO 6			
; LENGTH: 829			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-678-160A-6			
Query Match 100.0%; Score 4369; DB 5; Length 829;			

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 829; Conservative 0; DB 5; Length 829;

QY 1 MGLPRGLASLLQLQVWLCQAASEPCRAVFAEAEVLEAGGAEQBPQALGKVFVCGPG 60
DB 1 MGLPRGLASLLQLQVWLCQAASEPCRAVFAEAEVLEAGGAEQBPQALGKVFVCGPG 60

QY 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRIILRRHKRDVWVAPISVPENG 120
DB 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRIILRRHKRDVWVAPISVPENG 120

QY 121 KGPFFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREEIAK 180
DB 121 KGPFFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREEIAK 180

QY 181 YELFGHVSNGASVEDPMNISIIIVTDQNDHKPKFTQDTPRGSVLEGLPGTSVMQVTAT 240
DB 181 YELFGHVSNGASVEDPMNISIIIVTDQNDHKPKFTQDTPRGSVLEGLPGTSVMQVTAT 240

QY 241 DEDDAIYTVNGVAVSIHSGEPKDPHDLMTIHRSTGTISVISGLDRKVPYETLTIOA 300
DB 241 DEDDAIYTVNGVAVSIHSGEPKDPHDLMTIHRSTGTISVISGLDRKVPYETLTIOA 300

QY 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVQRLTVTDLDPNSP 360
DB 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVQRLTVTDLDPNSP 360

QY 361 AWRATYLMGDDGDHFTITTHPESNQILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
DB 361 AWRATYLMGDDGDHFTITTHPESNQILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420

QY 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480
DB 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480

QY 481 DPAGWLAMPDPSQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGTLTLID 540
DB 481 DPAGWLAMPDPSQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGTLTLID 540

QY 541 VNDHGPVPEPRQITICNOSPVRHVLNITDKLSPTSPPQAQLTDDSDIYWTAEVNEEGD 600
DB 541 VNDHGPVPEPRQITICNOSPVRHVLNITDKLSPTSPPQAQLTDDSDIYWTAEVNEEGD 600

QY 601 TVVLSLKKFLKQDITYDVHLSLSDHGNKEQLTVIRATVCDCHGVETCPGPKGGFPLPVL 660
DB 601 TVVLSLKKFLKQDITYDVHLSLSDHGNKEQLTVIRATVCDCHGVETCPGPKGGFPLPVL 660

QY 661 GAVLALLFLLLVLLVLRKKRKIKPEPLLLPEDDTRDNVFFYEGEGGEDDQDYITQLHR 720
DB 661 GAVLALLFLLLVLLVLRKKRKIKPEPLLLPEDDTRDNVFFYEGEGGEDDQDYITQLHR 720

QY 721 GLEARPEVLRNDVAPTIPTPMYRPRANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
DB 721 GLEARPEVLRNDVAPTIPTPMYRPRANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780

QY 781 DYEGSGSDAASLSLSTSSASDQDYDYLNEGWSRFPKCLADMYGGEDD 829
DB 781 DYEGSGSDAASLSLSTSSASDQDYDYLNEGWSRFPKCLADMYGGEDD 829

RESULT 13
US-10-723-860-2835
; Sequence 2835, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26

PRIOR APPLICATION NUMBER: 60/429,739
PRIOR FILING DATE: 2002-11-26
NUMBER OF SEQ ID NOS: 8393
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2835
LENGTH: 829
TYPE: PRT
ORGANISM: Homo sapiens
US-10-723-860-2835

Query Match 100.0%; Score 4369; DB 5; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLPRGLASLLQLQVWLCQAASEPCRAVFAEAEVLEAGGAEQBPQALGKVFVCGPG 60
DB 1 MGLPRGLASLLQLQVWLCQAASEPCRAVFAEAEVLEAGGAEQBPQALGKVFVCGPG 60

QY 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRIILRRHKRDVWVAPISVPENG 120
DB 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRIILRRHKRDVWVAPISVPENG 120

QY 121 KGPFFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREEIAK 180
DB 121 KGPFFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREEIAK 180

QY 181 YELFGHVSNGASVEDPMNISIIIVTDQNDHKPKFTQDTPRGSVLEGLPGTSVMQVTAT 240
DB 181 YELFGHVSNGASVEDPMNISIIIVTDQNDHKPKFTQDTPRGSVLEGLPGTSVMQVTAT 240

QY 241 DEDDAIYTVNGVAVSIHSGEPKDPHDLMTIHRSTGTISVISGLDRKVPYETLTIOA 300
DB 241 DEDDAIYTVNGVAVSIHSGEPKDPHDLMTIHRSTGTISVISGLDRKVPYETLTIOA 300

QY 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVQRLTVTDLDPNSP 360
DB 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVQRLTVTDLDPNSP 360

QY 361 AWRATYLMGDDGDHFTITTHPESNQILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
DB 361 AWRATYLMGDDGDHFTITTHPESNQILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420

QY 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480
DB 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480

QY 481 DPAGWLAMPDPSQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGTLTLID 540
DB 481 DPAGWLAMPDPSQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGTLTLID 540

QY 541 VNDHGPVPEPRQITICNOSPVRHVLNITDKLSPTSPPQAQLTDDSDIYWTAEVNEEGD 600
DB 541 VNDHGPVPEPRQITICNOSPVRHVLNITDKLSPTSPPQAQLTDDSDIYWTAEVNEEGD 600

QY 601 TVVLSLKKFLKQDITYDVHLSLSDHGNKEQLTVIRATVCDCHGVETCPGPKGGFPLPVL 660
DB 601 TVVLSLKKFLKQDITYDVHLSLSDHGNKEQLTVIRATVCDCHGVETCPGPKGGFPLPVL 660

QY 661 GAVLALLFLLLVLLVLRKKRKIKPEPLLLPEDDTRDNVFFYEGEGGEDDQDYITQLHR 720
DB 661 GAVLALLFLLLVLLVLRKKRKIKPEPLLLPEDDTRDNVFFYEGEGGEDDQDYITQLHR 720

QY 721 GLEARPEVLRNDVAPTIPTPMYRPRANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
DB 721 GLEARPEVLRNDVAPTIPTPMYRPRANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780

QY 781 DYEGSGSDAASLSLSTSSASDQDYDYLNEGWSRFPKCLADMYGGEDD 829
DB 781 DYEGSGSDAASLSLSTSSASDQDYDYLNEGWSRFPKCLADMYGGEDD 829

RESULT 14
US-10-696-639-39

```
; Sequence 39, Application US/10696639
; Publication No. US20050037439A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia Corporation
; APPLICANT: Bourner, Maureen J.
; TITLE OF INVENTION: DIFFERENTIALY EXPRESSED GENES INVOLVED IN CANCER, THE
; TITLE OF INVENTION: POLYPEPTIDES ENCODED THEREBY, AND METHODS OF USING THE SAME
; FILE REFERENCE: 01040/1
; CURRENT APPLICATION NUMBER: US/10/696,639
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: 60/422,176
; PRIOR FILING DATE: 2002-10-29
; NUMBER OF SEQ ID NOS: 3114
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 829
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-696-639-39

Query Match      100.0%; Score 4369; DB 5; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MGLPRGPLASLLLLQVLCWQCAASEPCRAVFRBAEVTLEAGGAEQEQALGKVFMGCPG 60
Db      1 MGLPRGPLASLLLLQVLCWQCAASEPCRAVFRBAEVTLEAGGAEQEQALGKVFMGCPG 60

Qy      61 QEPALFSTDNDFTVRNGETVOERSSLKERNPLKIPFSKRILRRHKRDVWVAPISVPENG 120
Db      61 QEPALFSTDNDFTVRNGETVOERSSLKERNPLKIPFSKRILRRHKRDVWVAPISVPENG 120

Qy      121 KGPPFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREBIAK 180
Db      121 KGPPFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREBIAK 180

Qy      121 KGPPFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREBIAK 180
Db      121 KGPPFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREBIAK 180

Qy      181 YELFGHVAENGASVEDPMNISIIVTQNDHKPKFTQDTFRGSLVLEGVLPGETSVMQVTAT 240
Db      181 YELFGHVAENGASVEDPMNISIIVTQNDHKPKFTQDTFRGSLVLEGVLPGETSVMQVTAT 240

Qy      241 DEDDAIYTYNGVAVSIHQSPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
Db      241 DEDDAIYTYNGVAVSIHQSPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300

Qy      241 DEDDAIYTYNGVAVSIHQSPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
Db      241 DEDDAIYTYNGVAVSIHQSPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300

Qy      721 GLEARPEVVLNRNDVAPTIIPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
```

```
Db      721 GLEARPEVVLNRNDVAPTIIPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
Qy      781 DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGDD 829
Db      781 DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGDD 829

RESULT 15
US-10-948-518-145
; Sequence 145, Application US/10948518
; Publication No. US2005006492A1
; GENERAL INFORMATION:
; APPLICANT: FREDERIC J. DESAUVAGE
; APPLICANT: GRETCHEN FRANTZ
; APPLICANT: KENNETH J. HILLAN
; APPLICANT: PAUL POLAKIS
; APPLICANT: ANDREW POLSON
; APPLICANT: VICTORIA SMITH
; APPLICANT: SUSAN D. SPENCER
; APPLICANT: THOMAS D. WU
; APPLICANT: ZEMIN ZHANG
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5026R1-US
; CURRENT APPLICATION NUMBER: US/10/948,518
; CURRENT FILING DATE: 2004-09-22
; PRIOR APPLICATION NUMBER: US/10/643,795
; PRIOR FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US 60/413,192
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US 60/419,008
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/426,847
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/484,959
; PRIOR FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 145
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-948-518-145

Query Match      100.0%; Score 4369; DB 5; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MGLPRGPLASLLLLQVLCWQCAASEPCRAVFRBAEVTLEAGGAEQEQALGKVFMGCPG 60
Db      1 MGLPRGPLASLLLLQVLCWQCAASEPCRAVFRBAEVTLEAGGAEQEQALGKVFMGCPG 60

Qy      61 QEPALFSTDNDFTVRNGETVOERSSLKERNPLKIPFSKRILRRHKRDVWVAPISVPENG 120
Db      61 QEPALFSTDNDFTVRNGETVOERSSLKERNPLKIPFSKRILRRHKRDVWVAPISVPENG 120

Qy      121 KGPPFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREBIAK 180
Db      121 KGPPFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREBIAK 180

Qy      181 YELFGHVAENGASVEDPMNISIIVTQNDHKPKFTQDTFRGSLVLEGVLPGETSVMQVTAT 240
Db      181 YELFGHVAENGASVEDPMNISIIVTQNDHKPKFTQDTFRGSLVLEGVLPGETSVMQVTAT 240

Qy      241 DEDDAIYTYNGVAVSIHQSPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
Db      241 DEDDAIYTYNGVAVSIHQSPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300

Qy      301 TMDGDGSTTTTAVAVVEILDANDNAPMDFDQKYEAHVPENAVGHVQRLTVTDLDPNSP 360
```

301	Db	 TMDGSGSTTTAVA VVELDANDNAPMFPDQKYEAHVPENAVGHVEQRLVTVDLDAPNSP	360
361	Qy	 AWRATYLIIMGDDGDHFTITTHPESNOQILITTRKGLDPEAKNOHTLVVEVTNEAPVFKL	420
361	Db	 AWRATYLIIMGDDGDHFTITTHPESNOQILITTRKGLDPEAKNOHTLVVEVTNEAPVFKL	420
421	Qy	 PTSTATIVVHVEDVNEAPVFVPPSKVVEVQSGIPTGSPVCVYTAEDDPKENQKISYRILR	480
421	Db	 PTSTATIVVHVEDVNEAPVFVPPSKVVEVQSGIPTGSPVCVYTAEDDPKENQKISYRILR	480
481	Qy	 DPAGWLAMPDSDGQVTA VGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLTLD	540
481	Db	 DPAGWLAMPDSDGQVTA VGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLTLD	540
541	Qy	 VNDHGPVPPEPROITICNOSPVRHVNLNITDKOLSPHTSPFQAQLTDDSDIYWTAEVNEEGD	600
541	Db	 VNDHGPVPPEPROITICNOSPVRHVNLNITDKOLSPHTSPFQAQLTDDSDIYWTAEVNEEGD	600
601	Qy	 TVVLSLKXFLKQDVTYDVHLSLSDHGKNEQLTVIRATVCDCHGHVETCPGPMWGGFLLPVL	660
601	Db	 TVVLSLKXFLKQDVTYDVHLSLSDHGKNEQLTVIRATVCDCHGHVETCPGPMWGGFLLPVL	660
661	Qy	 GAVLALLFLVLLVLLVLRKKIKIPEPLLPEDDTRDNVFFYGGEGGSEEDQDYDTQLHR	720
661	Db	 GAVLALLFLVLLVLLVLRKKIKIPEPLLPEDDTRDNVFFYGGEGGSEEDQDYDTQLHR	720
721	Qy	 GLEARPEVVLNRDVAPTIIPTPMYRPRPANPEIGNFIIENLKAANTDPTAPPYDTLLVF	780
721	Db	 GLEARPEVVLNRDVAPTIIPTPMYRPRPANPEIGNFIIENLKAANTDPTAPPYDTLLVF	780
781	Qy	 DYEAGSGDAASLSSITSSASDQDQDYDYLNEWGSRFKKLADMYGGGDD	829
781	Db	 DYEAGSGDAASLSSITSSASDQDQDYDYLNEWGSRFKKLADMYGGGDD	829

Search completed: December 12, 2005, 09:19:52
Job time : 171 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2005, 09:02:47 ; Search time 235 Seconds
(without alignments)

2488.864 Million cell updates/sec

Title: US-10-696-639-39

Perfect score: 4369

Sequence: 1 MGLPRGPLASLLLLQVWLQ.....NEWGSRFKKLADMYGGEDD 829

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4369	100.0	829	1 CADH3_HUMAN	P22223 homo sapien
2	3564.5	81.6	822	2 Q8BSL6_MOUSE	Q8bsl6 mus musculu
3	3560	81.5	821	2 Q8BRE1_MOUSE	Q8bre1 mus musculu
4	3550.5	81.3	822	1 CADH3_MOUSE	P10287 mus musculu
5	3433.5	78.6	753	2 Q80YV6_MOUSE	Q80yv6 mus musculu
6	2573.5	58.9	732	1 CADHK_CHICK	P33145 gallus gall
7	2420	55.4	882	1 CADH1_HUMAN	P12830 homo sapien
8	2400.5	55.1	882	2 Q6RBF2_BOVIN	Q6rf2 bos taurus
9	2400.5	54.9	901	2 Q9UII7_HUMAN	Q9uii7 homo sapien
10	2388	54.7	884	1 CADH1_MOUSE	P09803 mus musculu
11	2388	54.7	884	2 Q4KML8_MOUSE	Q4kml8 mus musculu
12	2388	54.7	886	1 CADH1_RAT	Q9r0t4 rattus norv
13	2387	54.6	882	2 Q5RAX1_PONPY	Q5rax1 pongo pygma
14	2377	54.4	884	2 Q6NTM0_XENLA	Q6ntm0 xenopus lae
15	2375	54.4	880	1 CADHP0_XENLA	P33148 xenopus lae
16	2371	54.3	884	1 CADH8_XENLA	P33152 xenopus lae
17	2323	53.2	431	1 CADH3_BOVIN	P19535 bos taurus
18	2274	52.0	887	1 CADH1_CHICK	P08641 gallus gall
19	2216.5	50.7	821	2 Q9UII8_HUMAN	Q9uii8 homo sapien
20	2193	50.3	872	1 CADH1_XENLA	P30344 xenopus lae
21	1893	43.3	864	2 Q90237_BRARE	Q90237 brachydanio
22	1869	42.8	776	2 Q4TX7_TETNG	Q4tx7 tetraodon n
23	1715	39.3	906	2 Q5RAX1_PONPY	Q5rax1 pongo pygma
24	1714	39.2	906	1 CADH2_HUMAN	P10222 homo sapien
25	1704.5	39.0	912	1 CADH2_CHICK	P10288 gallus gall
26	1702	39.0	906	2 Q8N173_HUMAN	Q8n173 homo sapien
27	1701	38.9	906	2 Q8BSI9_MOUSE	Q8bsi9 mus musculu
28	1701	38.9	906	2 Q6GU11_MOUSE	Q6gu11 mus musculu
29	1697.5	38.9	906	1 CADH2_MOUSE	P15116 mus musculu
30	1695	38.8	1101	2 Q4TX6_TETNG	Q4tx6 tetraodon n
31	1694.5	38.8	875	2 Q5RB18_PONPY	Q5rb18 pongo pygma

32	1694	38.8	913	1 CADH4_CHICK	P24503 gallus gall
33	1693	38.8	906	1 CADH2_RAT	Q92ly3 rattus norv
34	1691	38.7	877	1 CADH2_BOVIN	P19534 bos taurus
35	1691	38.5	893	1 CADH2_BRARE	Q90275 brachydanio
36	1659.5	38.0	876	2 Q4RVH8_TETNG	Q4rvh8 tetraodon n
37	1655	37.9	913	1 CADH4_MOUSE	P39038 mus musculu
38	1648.5	37.7	916	2 Q9BZ05_HUMAN	Q9bz05 homo sapien
39	1638	37.5	824	2 Q8NB64_HUMAN	Q8nb64 homo sapien
40	1621.5	37.1	916	1 CADH4_HUMAN	P55283 homo sapien
41	1616.5	37.0	906	1 CADHO_XENLA	P33147 xenopus lae
42	1609	36.8	905	1 CADHN_XENLA	P20310 xenopus lae
43	1606	36.8	933	2 Q4UII6_BRARE	Q4uii6 brachydanio
44	1592	36.4	908	2 Q4SS36_TETNG	Q4ss36 tetraodon n
45	1578.5	36.1	922	2 P79883_9PIPI	P79883 xenopus. ma

ALIGNMENTS

RESULT 1
CADH3_HUMAN
ID CADH3_HUMAN STANDARD; PRT; 829 AA.
AC P22223;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Cadherin-3 precursor (Placental-cadherin) (P-cadherin).
GN Name=CDH3; Synonyms=CDHP;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=90009051; PubMed=2793940; DOI=10.1083/jcb.109.4.1787;
RA Shimoyama Y., Yoshida T., Terada M., Shimamoto Y., Abe O.,
RA Hirohashi S.;
RT "Molecular cloning of a human Ca2+-dependent cell-cell adhesion
molecule homologous to mouse placental cadherin: its low expression in
human placental tissues";
RL J. Cell Biol. 109:1787-1794 (1989).
[2]
RN
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
TX TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klauser R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Aramori P.,
Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
proc. mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[3]
RN
RP NUCLEOTIDE SEQUENCE OF 1-15.
TX TISSUE=Fetal brain;
RX MEDLINE=99111163; PubMed=9815605;
RA Jarraud D.F., Paul R., Van Bokhoven A., Nguyen S.H., Bova G.S.,
Wheelock M.J., Johnson K.R., Schaiken J., Bussemakers M., Isaacs W.B.;
RT "P-cadherin is a basal cell-specific epithelial marker that is not

RT expressed in prostate cancer.";
RL Clin. Cancer Res. 3:2121-2128(1997).
RN [4].
RP TISSUE SPECIFICITY.
RX MEDLINE=89195140; PubMed=2702654;
RA Shimoyama Y., Hirohashi S., Hirano S., Noguchi M., Shimoseato Y.,
RA Takeichi M., Abe O.;
RT "Cadherin cell-adhesion molecules in human epithelial tissues and
RT carcinomas";
RL Cancer Res. 49:2128-2133(1989).
RN [5].
RN DISEASE.
RX MEDLINE=21470328; PubMed=11544476; DOI=10.1038/ng716;
RP Sprecher E., Bergman R., Richard G., Lurie R., Shalev S.,
RA Petronius D., Shalata A., Ambinder Y., Leibin R., Periman I., Cohen N.,
RA Szargel R.;
RT "Hypotrichosis with juvenile macular dystrophy is caused by a mutation
RT in CDH3, encoding P-cadherin";
RL Nat. Genet. 29:134-136(2001).
RN [6].
RN VARIANT HJMD HIS-503.
RX MEDLINE=22336734; PubMed=12445216;
RX DOI=10.1046/j.1523-1747.2002.19528.x;
RA Indelman M., Bergman R., Lurie R., Richard G., Miller B.,
RA Petronius D., Ciubutaro D., Leibin R., Sprecher E.;
RT "A missense mutation in CDH3, encoding P-cadherin, causes
RT hypotrichosis with juvenile macular dystrophy.";
RL J. Invest. Dermatol. 119:1210-1213(2002).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in some normal epithelial tissues
CC and in some carcinoma cell lines.
CC -!- DISEASE: Defects in CDH3 are the cause of hypotrichosis with
CC juvenile macular dystrophy (HJMD) [MIM:601553]. HJMD is a rare
CC autosomal recessive disorder characterized by early hair loss
CC heralding severe degenerative changes of the retinal macula and
CC culminating in blindness during the second to third decade of
CC life.
CC -!- SIMILARITY: Contains 5 cadherin domains.
CC -!- DATABASE: NAME=Mutations of the CDH3 gene;
CC NOTE=Retina International's Scientific Newsletter;
CC WWW="http://www.retina-international.com/sci-news/cdh3mut.htm".
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X63629; CAA45177.1; -; mRNA.
DR EMBL; BC041846; AAA41846.1; -; mRNA.
DR EMBL; X95824; CAA65093.1; -; Genomic_DNA.
DR PIR; A33659; IJHUCP.
DR HSSP; P09803; 117W.
DR SMR; P22223; 749-824.
DR Ensembl; ENSG00000602038; Homo sapiens.
DR HGNC; HGNC:1762; CDH3.
DR MIM; 114021; -.
DR MIM; 601553; -.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR00126; Cadherin.
DR InterPro; IPR000233; Cadherin_C-term.
DR Pfam; PF00028; Cadherin; 5.
DR Pfam; PF01049; Cadherin_C; 1.
DR PRINTS; PR0205; CADHERIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS0268; CADHERIN_2; 5.
KW Calcium; Cell adhesion; Disease mutation; Glycoprotein; Repeat;
KW Sensory transduction; Signal; Transmembrane; Vision.

FT	SIGNAL	1	24	Potential.
FT	PROPEP	25	107	
FT	CHAIN	108	829	Cadherin-3.
FT	TOPO_DOM	108	654	Extracellular (Potential).
FT	TRANSMEM	655	677	Potential.
FT	TOPO_DOM	678	829	Cytoplasmic (Potential).
FT	DOMAIN	108	215	Cadherin 1.
FT	DOMAIN	216	328	Cadherin 2.
FT	DOMAIN	329	440	Cadherin 3.
FT	DOMAIN	441	546	Cadherin 4.
FT	DOMAIN	547	650	Cadherin 5.
FT	COMPTIAS	785	800	Ser-rich.
FT	CARBOHYD	200	200	N-linked (GLCNAC. . .) (Potential).
FT	CARBOHYD	566	566	N-linked (GLCNAC. . .) (Potential).
FT	VARIANT	503	503	R -> H (in HJMD). /FTId=VAR_015422.
FT	CONFLICT	237	237	V -> M (in Ref. 2).
FT	CONFLICT	563	563	H -> Q (in Ref. 2).
FT	SEQUENCE	829	AA; 91427	MW; E503CFFFE5D981F1 CRC64;
Qy	Query Match	100.0%	Score 4369; DB 1; Length 829;	
Db	Best Local Similarity	100.0%;	Pred. No. 2.3e-264;	
Qy	Matches 829; Conservative	0; Mismatches	0; Indels	0; Gaps
Db	1	MGLPRGLASLLLLQVLCQAASEPCRAVFREAEVTL	LEAGAEQEPGQALGVFMGCPG	60
Qy	1	MGLPRGLASLLLLQVLCQAASEPCRAVFREAEVTL	LEAGAEQEPGQALGVFMGCPG	60
Db	61	QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIP	PSKRILRRHKRDWVAPISVPENG	120
Qy	61	QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIP	PSKRILRRHKRDWVAPISVPENG	120
Db	121	KGPPORLNOLKSNKORDTKIFYISITGPGADSP	EGVFAVEKETGMLLNKPLDREIAK	180
Qy	121	KGPPORLNOLKSNKORDTKIFYISITGPGADSP	EGVFAVEKETGMLLNKPLDREIAK	180
Db	181	YELFGHAVSENGASVEDPMNISIIVTDQNDHKPK	FTQDTPRGSLGVLPGTSMQVTAT	240
Qy	181	YELFGHAVSENGASVEDPMNISIIVTDQNDHKPK	FTQDTPRGSLGVLPGTSMQVTAT	240
Db	241	DEDDAIYTYNGVAVSIHSGEPKDPHDLMTIHR	STGTISVISGDLREKVPYTLTQA	300
Qy	241	DEDDAIYTYNGVAVSIHSGEPKDPHDLMTIHR	STGTISVISGDLREKVPYTLTQA	300
Db	301	TDMDGSGSTTAVAVVEILLDANDNAPMDFPQK	YEAHVPENAVGHVORLTVTDLDPNSP	360
Qy	301	TDMDGSGSTTAVAVVEILLDANDNAPMDFPQK	YEAHVPENAVGHVORLTVTDLDPNSP	360
Db	361	AMRATYLLMGDDGDHFTITTHPESNQGLITTR	KGLDFAKNQHTLYVEVTNEAPFVLKL	420
Qy	361	AMRATYLLMGDDGDHFTITTHPESNQGLITTR	KGLDFAKNQHTLYVEVTNEAPFVLKL	420
Db	421	PTSTATIVHVEDVNEAPVFPVPPSKVVEQEGIP	TGEPVCVYTAEDDPKKNQKISYILR	480
Qy	421	PTSTATIVHVEDVNEAPVFPVPPSKVVEQEGIP	TGEPVCVYTAEDDPKKNQKISYILR	480
Db	481	DPAGWLAMPDSCQVTAAGTLDREDEQFVNNI	YEVWVLAMWDNGSPPTTGTGTLTLLID	540
Qy	481	DPAGWLAMPDSCQVTAAGTLDREDEQFVNNI	YEVWVLAMWDNGSPPTTGTGTLTLLID	540
Db	541	VNDHGPVPEPRQITICNQSPVRHVLNITDKD	LSPTSPFOAQLTDDSDIYWTAEVNEEGD	600
Qy	541	VNDHGPVPEPRQITICNQSPVRHVLNITDKD	LSPTSPFOAQLTDDSDIYWTAEVNEEGD	600
Db	601	TVVLSLKKFLKQDTPVHLSLSDHGNKQQLTV	IRATVCDCHGHVETCPGPKWKGFIPLVL	660
Qy	601	TVVLSLKKFLKQDTPVHLSLSDHGNKQQLTV	IRATVCDCHGHVETCPGPKWKGFIPLVL	660
Db	661	GAVLALLFLLLVLLVLLVLRKKKIKPELLLP	EDTRDNVYVYGGEGGEEQDDYDITQHR	720
Qy	661	GAVLALLFLLLVLLVLLVLRKKKIKPELLLP	EDTRDNVYVYGGEGGEEQDDYDITQHR	720
Db	721	GLEARPEVLRNDVAPTIITPTMYRPRPANP	DEIGNFIENLKAANTDPTAPPYDTLLVF	780
Qy	721	GLEARPEVLRNDVAPTIITPTMYRPRPANP	DEIGNFIENLKAANTDPTAPPYDTLLVF	780

```

Db      721 GLEAREVVLNRDVAFTIPTMYRPRPANPDBEIGNFIENLKAANTDPTAPPYDTLLVF 780
QY      781 DYEAGSDAASLSSTSSADODQDYDLNWSRFFKLADMYGGGEDD 829
Db      781 DYEAGSDAASLSSTSSADODQDYDLNWSRFFKLADMYGGGEDD 829

RESULT 2
Q8BSL6 MOUSE
ID Q8BSL6 MOUSE PRELIMINARY; PRT; 822 AA.
AC Q8BSL6
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length enriched
DE library, clone:5930439E13 product:cadherin 3, full insert sequence
DE (Cadherin 3).
DE Names:cdh3;
GN Names:cdh3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayaehizaki Y.;
RT "high-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=21085650; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gijobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bernaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayaehizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayaehizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [5]

```

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Tahii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayaehizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Haezizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa T., Kasukawa T.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayaehizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Stapleton M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Locquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RG NIH MGC Project;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK031265; BAC27327.1; -; mRNA.
DR EMBL; BC098459; AAH98459.1; -; mRNA.
DR HSSP; P09803; 117X.
DR Ensemble; ENSMUSG00000061048; Mus musculus.
DR MGI; MGI:88356; Cdh3.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; IDA.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0016337; P:cell-cell adhesion; IMP.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C term.
DR Pfam; PF001028; Cadherin; 5.
DR Pfam; PF01049; Cadherin; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 3.

```

DR PROSITE; PSS0268; CADHERIN_2; 4.
KW Calcium.
SQ SEQUENCE 822 AA; 90612 MW; 7C653D58210A595A CRC64;

Query Match 81.6%; Score 3564.5; DB 2; Length 822;
Best Local Similarity 81.9%; Pred. No. 4.6e-214;
Matches 681; Conservative 56; Mismatches 82; Indels 13; Gaps 5;

QY 1 MGLPRGFLA-SLLLLQVCWLQCAASEPCRAVP-REAETVLEAGGAEQEPGQALGVFMGC 58
DB 1 MELLSGPHAFLLLLQVCWLSVSVSEPRAGFIGEAGVTLEVGTDLPSPQVLGKVALAG 60
QY 59 PGQEPALFSTDNDDFTVRNGETVQGRSLKERNPLKIPPSKRILRRHKRDWVVAIPISVE 118
DB 61 QGMHHA----DNGDIIMLTGCTVQGGKDAHSE-----PPTRILRRRRKREWWMPPIFVPE 110

QY 119 NGKPPFORLNQLSKNDKRDTKIFYSITGPGADSPPEGVFAVEKETGHWLLNKLPLDREI 178
DB 111 NGKPPFORLNQLSKNDKRGTKIFYSITGPGADSPPEGVFTIEKSGWLLHMPDLDRDKI 170

QY 179 AKYELFGHAYSENGASVEDPMNISIIIVTDQNDHKPKFTQDTPRGSVLGVLPGTSMQVOT 238
DB 171 VKIELYGHAYSENGASVEEPMNISIIIVTDQNDKRPKFTQDTPRGSVLGVLPGTSMQVOT 230

QY 239 ATDEDDAIYTVNGVAYSIIHQEPKPDHLMFTTHRSSTGTISVISGLDREKVPYTLITI 298
DB 231 ATDEDDAVNTYNGVAYSIIHQEPKPDHLMFTTHKSTGTISVISGLDREKVPYRLTV 290

QY 299 QATDMGSGSTTAVAVVEILDANDNAMPFPQKYEAVHPENAVGHEVORLTVTDLDAPN 358
DB 291 QATDMGSGSTTAAVAVVQIILDANDNAPEFPQKYEAVHPENAVGHEVORLTVTDLDVFN 350

QY 359 SPWARTVILMGDDGDHFTTHTPESNOGLITTRKGLDFAKNQHTLYVEVTNEAPFVL 418
DB 351 SPWARTVHI VGGDDGDHFTTHTPETNQGLVTLTKGLDFAQDQHTLYVEVTNEAPAV 410

QY 419 KLPTSTATIVVHVDVNEAPVFPVPSKVVEVQEGIPTGEPVCVYTAEDPKENQKISYRI 478
DB 411 KLPTATATVVHVVDVNEAPVFPVPSKVIEAQEGISIGELVCITYAQDPDKEDQKISYTI 470

QY 479 LRDPAGMLAMPDSQGVTAAGTLDRDQFVRNNIYEVWVLAMDNQSPPTTGTLILLTL 538
DB 471 SRDPANMLAVPDSQGIITAAAGILDRDQFVKNNYEVWVLATDSGNPPTTGTLILLTL 530

QY 539 IDVNDHGVPSPEPRTICNQSPVRHLVNTDKLSPTSPPOAQLTDSDIYVTAEVNEE 598
DB 531 TDINDHGPPEPRQIIICNQSPVPQVLNITDKLSFNSPPOAQLTDSDIYVMAEVSEK 590

QY 599 GDTVLSLKKFLKQDYTVHLSLSHGNKEQLTVIRATVCDCHGV-ETCPGPWKGFFIL 657
DB 591 GDTVALSLKKFLKQDYTDYDLHLSLSDHGNREQLTMIRATVCDCHGVFNDPWPWKGFFIL 650

QY 658 PVLGAVLALLFLLVLLVLRKKRKIKPSPLLLPEDDTRDNVFPYEGEGGEDQDYDITQ 717
DB 651 PILGAVLALLTLLALLVLRKKRKVKSPLLLPEDDTRDNVFPYEGEGGEDQDYDITQ 710

QY 718 LHRGLEARPEVLRNDVAPTIIPTPMYPRPANPDEIGNFIENLKAANTDPTAPPYDTL 777
DB 711 LHRGLEARPEVLRNDVVPFTIPTPMYPRPANPDEIGNFIENLKAANTDPTAPPYDSL 770

QY 778 LVFDYEGSGDAASLSLSTSSASDQDQDYVLNENSGSRFPKLLADMYGGGEDD 829
DB 771 LVFDYEGSGDAASLSLSTSSASDQDQDYVLNENSGSRFPKLLADMYGGGEDD 822

RESULT 3
Q8BRE1_MOUSE
ID Q8BRE1_MOUSE PRELIMINARY; PRT; 821 AA.
AC Q8BRE1_MOUSE
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 9.5 days embryo parthenogenote cDNA, RIKEN full-length

Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishi Y., Itoh M., Kagawa I., Kasukawa T., Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito R., Saichon H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akaira S., Takada Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.; Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK045041; BAC32194.1; -, mRNA.
DR HSSP; P09803; 117X.
DR MGI; MGI:88356; Cdh3.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0016337; P:cell-cell adhesion; IMP.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; Cadherin; 5.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS00268; CADHERIN_2; 4.
KW Calcium.
SQ SEQUENCE 821 AA; 90513 MW; 168B56C5FB5CCED0 CRC64;

Query Match 81.5%; Score 3560; DB 2; Length 821;
Best Local Similarity 81.6%; Pred. No. 8.9e-214;
Matches 679; Conservative 56; Mismatches 83; Indels 14; Gaps 5;

QY 1 MGIPRGLA-SLLQLQVCMLOCAASPCRAVF-REAEVTLAAGASQEPQALGKVFMC 58
DB 1 MEULSGPHALLLLQLQVCMVVSVSPYRAGFTGEAGVILEVETGDLQSPQLGKALAG- 59

QY 59 PGQEPALFSDNDFTVRNGETVQERRSLKERNPLKIFPSKRLRRHKRDWVAVPSPE 118
DB 60 -----QGMHADNGDIIMLTGTVQGGKAMHS-----PPTRLRRKREWMVPPFVPE 109

QY 119 NGKGPPQRLNQLKSKNRDRTKIFYSITGPADSPPEGVFAVKEFTGWLNLKPLDRBEI 178
DB 110 NGKGPPQRLNQLKSKNRDRTKIFYSITGPADSPPEGVFTIEKSGWLLHMLPLDREKI 169

QY 179 AKYELCHAVSENGASVEQPMNISIIVTDNDHKPKFTQDTFGSLVGLVPGTSVMQVT 238
DB 170 VKYELCHAVSENGASVEQPMNISIIVTDQNDKPKFTQDTFGSLVGLVPGTSVMQVT 229

QY 239 ATDEDDAIVTYNGWVAYSISQBPDPKPHDLMTFTHRSTGTSIVISSGLDREKVPYTLTI 298
DB 230 ATDEDDAVNTYNGWVAYSISQBPKEPHDLMTFTHKSTGTSIVISSGLDREKVPYRLTV 289

QY 299 QATDMGDSGTTTAVAVEILDANDNAPMFPDQKYEAHVPENAVGHEVQRLTVDLDAPN 358
DB 290 QATDMGDSGTTTAAEAVVQILDANDNAPPEPQKYEAWVPENEVEHQRLTVDLDVFN 349

QY 359 SPAWRATYLIMGDDGDFHTIHTHPSNQIGILTRKGLDPEAKQHTLVYEVNTEAPFVL 418
DB 350 SPAWRATYHIVGGDGDGHFTIHTHPETNQGLVTKKGLDPEAQDQHTLVYEVNTEAPFAV 409

QY 419 KLPTSTATIVHVVEDVNEAPFVPPSKVVEVQSGIPTGPEVCVYTAEDPDKENOKISYRI 478
DB 410 KLPTATATVHVVDVNEAPFVPPSKVIEAQSGISIGELVCIYTAQDPDKEDQKISYTI 469

QY 479 LRDPAGWLAMPDPSGQVAVTGLDREDEQFVRNNIYEVWVLMNDGSPPTTGTGTLTLL 538
DB 470 SRDPANWLAVDPSGQITAAIGILDREDEQFVKNNVYEVWVLMATDSGNPPTTGTGTLTLL 529

QY 539 IDVNDHGVPPEPQITICNOSPVRVHLNITDKLSPTSPFOQLTDDSDIYKTAEVNEE 598
DB 530 TDINDHGPIPEPQIIICNOSPVPQVQLNITDKLSNPSFPQQLTHDSDIYKMAEVSEK 589

QY 599 GDTVVLSLKKFLKQTYDYVHLSLSHGHNKEQLTVIRATVCDCHGHV-ETCPGPKGGFLL 657
DB 590 GSTVALSLKKFLKQTYDYHLSLSHGHNKEQLTMIRATVCDCHGQVFNDCPRPWKGGFLL 649

QY 658 PVLGAVLALLFLLLVLLVLLVRKKRIKEPLLLPEDDTRDNVFFYGEGBEDQDYDITQ 717
DB 650 PILGAVLALLTLLALLLLVLRKKRKEPLLLPEDDTRDNVFFYGEGBEDQDYDITQ 709

QY 718 LHRGLEARPEVVLNRDVAPTIITPMYRPRNPANDEIGNFIENLKAANTDTPAPPYDTL 777
DB 710 LHRGLEARPEVVLNRDVAPTIITPMYRPRNPANDEIGNFIENLKAANTDTPAPPYDSL 769

QY 778 LVFDYEGSGSDAASLSLTSASDQDQDYDYLNWGSRFCKLADMYGGGEDD 829
DB 770 LVFDYEGSGSDAASLSLTSASDQDQDYDYLNWGSRFCKLADMYGGGEDD 821

RESULT 4
CADH3_MOUSE
ID CADH3_MOUSE STANDARD; PRT; 822 AA.
AC P10287; Q61465;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Cadherin-3 precursor (Placental-cadherin) (P-cadherin).
GN Name=Cdh3; Synonyms=Cdhp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=88111554; PubMed=3428270;
RA Nose A., Nagafuchi A., Takeichi M.;
RT "Isolation of placental cadherin cDNA: identification of a novel gene family of cell-cell adhesion molecules.";
RL EMBO J. 6:3655-3661(1987).
RN [2]
RP PARTIAL NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Embryo;
RX MEDLINE=91360343; PubMed=1886768;
RA Hatta M., Miyatani S., Copeland N.G., Gilbert D.J., Jenkins N.A., Takeichi M.;
RT "Genomic organization and chromosomal mapping of the mouse P-cadherin gene.";
RL Nucleic Acids Res. 19:4437-4441(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-55.
RX MEDLINE=93294853; PubMed=8515462;
RA Faraldo M.L., Cano A.;
RT "The 5' flanking sequences of the mouse P-cadherin gene. Homologies to 5' sequences of the E-cadherin gene and identification of a first 215 base-pair intron.";
RL J. Mol. Biol. 231:935-941(1993).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 1-55.
RC STRAIN=C57BL/6;
RA Hatta M., Takeichi M.;
RT "Complex cell type-specific transcriptional regulation by the promoter and an intron of the mouse P-cadherin gene.";
RL Dev. Growth Differ. 36:509-519(1994).
RN [5]
RP DEVELOPMENTAL STAGE.
RC STRAIN=C57BL/6; TISSUE=Testis;
RX MEDLINE=97033837; PubMed=8879495;
RA Munro S.B., Blaschuk O.W.;
RT "A comprehensive survey of the cadherins expressed in the testes of fetal, immature, and adult mice utilizing the polymerase chain reaction.";
RL Biol. Reprod. 55:822-827(1996).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC052189; AAH52189.1; -; mRNA.
DR HSSP; P09803; 117X.
DR MGI; MGI:88256; Cdh3.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0016337; P:cell-cell adhesion; IMP.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; Cadherin; 5.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS00268; CADHERIN_2; 4.
KW Calcium.
FT NON_TER.
SQ SEQUENCE 753 AA; 83300 MW; 414AFB4059CE94D2 CRC64;

Query Match 78.6%; Score 3433.5; DB 2; Length 753;
Best Local Similarity 87.1%; Pred. No. 6.4e-206;
Matches 639; Conservative 51; Mismatches 43; Indels 1; Gaps 1;

QY 97 PSKRILRRHRKDWVAPISVPENKGGPFPPORLNQKSNKDRDKIFYSITGPGADSPPEG 156
DB 20 PPRILRRKREWMPIFPVENKGGPFPPORLNQKSNKDRDKIFYSITGPGADSPPEG 79

QY 157 VFAVEKETGWLKLPDLREIEIAKYELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFT 216
DB 80 VFTIEKSGWLLHMLPDLREIKVYELYGHAVSENGASVEEPMNISIIIVTDQNDHKPKFT 139

QY 217 QDTFRGSLGVLPSTSMQVATDEDDALITYNGVAVSIHQEPKDPHDLMTTHRT 276
DB 140 QDTFRGSLGVLPSTSMQVATDEDDAVNTYNGVAVSIHQEPKDPHDLMTTHRT 199

QY 277 GTISVTSGLDRKVPREYTLTIQATDMGDSSTTAVVVEILDANDNAPMFPDQKYEAH 336
DB 200 GTISVTSGLDRKVPREYTLVQATDMGDSSTTAEAVVQILDANDNAPMFPDQKYEAW 259

QY 337 VPENAVGHEVQRLTVTDLDAPNSPAWRATYILMGDGDGDFHTTTHPESNQGLTTRKGL 396
DB 260 VPENAVGHEVQRLTVTDLDVPNSPAWRATYHIVGGDDGDFHTTTHPESNQGLTTRKGL 319

QY 397 DFRAKQHTLYVETNEAPVLKLPSTATIVHVEDVNEAPVFPVPPSKVEVQEGIPTG 456
DB 320 DFEAQDQHTLYVETNEAPVAVKLPATATATVYVHVKDVNEAPVFPVPPSKVIEAQEGISIG 379

QY 457 EPVCVYTAEDPDKENOKISVRIILRDPAGVLAMPDPSQVTAICTLQDREDEQFVRNNIYEV 516
DB 380 ELVCITYTAQPDKEQDKISVTISRDPAWLAVDPSGQITAAIGLREDEQFVKNVYEV 439

QY 517 MVLAMQNSPPTGTGTGLLTLLDVNDHGVPPFPQITICNQSPFVRHVLNITDKLSPHIT 576
DB 440 MVLATSGNPPTGTGTGLLTLLTDINDHGPPEPQIICNQSPVQVQLNITDKLSPPNS 499

QY 577 SPFQAQLTDSDIYWTAEVNEEGDVTLSLKKFLKQDVTYVHLSLSDHGNKEQLTVIRAT 636
DB 500 SPFQAQLTDSDIYWMAEVSEKGDVTALSLLKKFLKQDVTYVHLSLSDHGNKEQLTVIRAT 559

QY 637 VCDCHGV-ETCPGPKWGGFIPVLGAVLALLFLLVLLVLLVKKKIKPEPLLPEDDT 695
DB 560 VCDCHGVQFNDGCPKPKWGGFIPVLGAVLALLFLLVLLVLLVKKKIKPEPLLPEDDT 619
QY 696 DNVFYEGEGGEDQDITQLHRLGLARPEVLRNDVAPTIIPTMYRPRANPDEIG 755

DB 620 DNVFYEGEGGEDQDITQLHRLGLARPEVLRNDVAPTIIPTMYRPRANPDEIG 679
QY 756 NFIIENLKAANTDPTAPPYDTLLVPDYEGSGDAASLSLTSSASDQDDYDYLNWGSR 815
DB 680 NFIIENLKAANTDPTAPPYDTLLVPDYEGSGDAASLSLTSSASDQDDYDYLNWGSR 739
QY 816 FKLADMYGGEDD 829
DB 740 FKLADMYGGEDD 753

RESULT 6
CADHK_CHICK
ID CADHK_CHICK STANDARD; PRT; 732 AA.
AC P33145;
DC 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE B-cadherin precursor (K-CAM protein) (Fragment).
GN Name=K-CAM;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92107987; PubMed=1763068;
RA Sorkin B.C., Gallin W.J., Edelman G.M., Cunningham B.A.;
RT "Genes for two calcium-dependent cell adhesion molecules have similar
RT structures and are arranged in tandem in the chicken genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:11545-11549 (1991).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 7-732.
RC TISSUE=Embryonic brain;
RX MEDLINE=91225083; PubMed=2026653; DOI=10.1083/jcb.113.4.893;
RA Napolitano E.W., Venstrom K., Wheeler E.F., Reichardt L.F.;
RT "Molecular cloning and characterization of B-cadherin, a novel chick
RT cadherin.";
RL J. Cell Biol. 113:893-905 (1991).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types. B-cadherin may have important
CC functions in neurogenesis, in at least some epithelia, and in
CC embryogenesis.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in a wide variety of tissues.
CC -!- SIMILARITY: Contains 5 cadherin domains.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; M81894; AAA48929.1; -; Genomic_DNA.
CC EMBL; X58518; CAA41408.1; -; mRNA.
CC PIR; A41634; IJCHCB.
CC HSSP; P09803; 117X.
CC SMR; P33145; 7-224, 651-726.
CC InterPro; IPR002126; Cadherin.
CC InterPro; IPR000233; Cadherin_C_term.
CC Pfam; PF00028; Cadherin; 5.
CC Pfam; PF01049; Cadherin_C; 1.
CC PRINTS; PR00205; CADHERIN.
CC SMART; SM00112; CA; 4.
CC PROSITE; PS00232; CADHERIN_1; 3.
CC PROSITE; PS00268; CADHERIN_2; 5.
KW Calcium; Cell adhesion; Glycoprotein; Repeat; Transmembrane.
FT PROPEP <1 6 Potential.
FT CHAIN 7 732 B-cadherin.


```
FT TOPO_DOM 6 554 Extracellular (Potential).
FT TRANSMEM 555 580 Potential.
FT TOPO_DOM 581 732 Cytoplasmic (Potential).
FT DOMAIN 114 114 Cadherin 1.
FT DOMAIN 115 227 Cadherin 2.
FT DOMAIN 228 339 Cadherin 3.
FT DOMAIN 340 443 Cadherin 4.
FT DOMAIN 444 554 Cadherin 5.
FT COMPBIAS 689 702 Ser-rich.
FT CARBOHYD 137 137 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 410 410 N-linked (GlcNAc. .) (Potential).
FT CONFLICT 414 414 M -> V (in Ref. 2).
FT NON_TER 1
SQ SEQUENCE 732 AA; 80613 MW; 091D59A6A16CBD45 CRC64;

Query Match 58.9%; Score 2573.5; DB 1; Length 732;
Best Local Similarity 66.2%; Pred. No. 3.6e-152;
Matches 485; Conservative 103; Mismatches 138; Indels 7; Gaps 5;

Qy 102 LRRHKRWVAVIPSPENGKGFPPQRLNQLKSNKDRDTKIFYSITGPGADSPPRGFAVE 161
Db 1 LRRQKRWVPIPKVPENRGFPFNVLQIKSNDRDAKIFYSITGGADAPPEIGITIE 60

Qy 162 KETGLLLNKPDLREEIAKELFGLHVSNGASVEDPMNISIIIVTDQNDHKPKFTQDTR 221
Db 61 KETGMKVKTPQLDREHINKYHLVSHVSENGKPEVEPMELIIVTVDQNDKNKQFTQSVFR 120

Qy 222 GSVLEGLVPGTSMQVATDEDDAIYTYNGVVAISHSQPKDPHDLMTFIHRSTGISV 281
Db 121 GSVPEGALPGTSMVRNATDADDVETVNGVIAYSILSQEPREPHMFTVNRATGTLV 180

Qy 282 ISSGLDREKPEYTLTQATDMGDSGTTTAVAVVEILDANDNAPMFDPOKYEAAHPENA 341
Db 181 IASGLDRERVREYILTQWADLDQGLTTTALAVIEITVDNDNAPEDFPKTYEAAVENE 240

Qy 342 VGHEVQRLVTDLPAPNSPAWRATYLYINGSDGDGHFTTTHPESNOGILTRKGLDPEAK 401
Db 241 AELEVARLATDLDEPHTPAWRAVYSIVRGNEGGAFTITDPASNEGVLRKAKGLDYEA 300

Qy 402 NQHTLYVEVTEARFVLKLPSTATIVVHVEDUNEAFVPPSPKVVVEQGIPTGEPCV 461
Db 301 RQFVLHVAVNEARFAIKLPTATATVMVSEVDNEAFVDPDLRLAQVPEDVPLGQPLAS 360

Qy 462 YTAEDDPK-ENQKLSYRILTRDPAGWLAMPDPSGOVTVAGTLDREDOFVRNNIYEVMLA 520
Db 361 YTAQDPRAQQRIKYVMSGDPAGWLAVHPENGLITAREQLDRE-SPTKNSITWAVILA 419

Qy 521 MDNGSPPTGTGTLTLLTLDVNDHGPVPEPPROITCNQSPVVRHVLNITDKDLSPTSPFQ 580
Db 420 VDDGLPPATGATGTLTLLTLDVNDHGPPEPRDVIICNRSVPVQVLITTDRLDLPNTGPF 479

Qy 581 AQLTDDSDIYTAEVNEBGDTVVLSLKFLKQDTYDVHLSDSHGKNEQLTVIRATVDCD 640
Db 480 AELSHGSGDSWAVEGVNGSDTVALWLTEPLEQNLYSVYLRFLDRQGDQDQVTVIRAQV 539

Qy 641 HGHVETC---PGPWKG-GFILLPVLAGVALLFLALLVLLVLRKKRIKEPLLDPEDTRD 696
Db 540 QGRVESCAQKPRVDTGPIVLAVLAGVALLGLVLLVLLVLRKKVKEPLLDPEDTRD 599

Qy 697 NVFYGGEGGEGEDQDYDITQHRGLEARPEVRLRNDVAPTIITPMYRPRPAPNDPEIGN 756
Db 600 NIFYGGEGGEGEDQDYDLSQLHGLDARPEVI-RNDVAPPLMAAPQYRPRPAPNDPEIGN 658

Qy 757 FIENLKAANTDPTAPPYDITLLVFDYEGSGDAAASLSLTSSASDQDQDYDLNEWGRF 816
Db 659 FIDENLKAANTDPTAPPYDITLLVFDYEGGGSEATSLSLNLSASDQDQDYDLNEWGRF 718

Qy 817 KKLADMYGGGEDD 829
Db 719 KKLAELYGGGEDE 731
```

RESULT 7

```
CADH1_HUMAN
ID CADH1_HUMAN STANDARD; PRT; 882 AA.
AC P12830; Q13799; Q14216; Q15855; Q16194; Q4PJ14;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Epithelial-cadherin precursor (E-cadherin) (Uvomorulin) (Cadherin-1) (CAM 120/80).
GN Name=CDH1; Synonyms=CDHE, UVO;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Pancreas;
RX MEDLINE=93211394; PubMed=8459805;
RA Bussemakers M.J.G., Mees S.G.M., van Bokhoven A., Debruyne F.M.J.,
RA Schaiken J.A.;
RT "Molecular cloning and characterization of the human E-cadherin
RT cDNA.";
RL Mol. Biol. Rep. 17:123-128 (1993).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA].
RA Kelker W., Warda A., Oda T., Hirohashi S., Kemler R., Birchmeier W.;
RT "Sequence of human E-cadherin cDNA.";
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Liver;
RX MEDLINE=94242050; PubMed=8185635;
RA Rimm D.L., Morrow J.S.;
RT "Molecular cloning of human E-cadherin suggests a novel subdivision of
RT the cadherin superfamily.";
RL Biochem. Biophys. Res. Commun. 200:1754-1761 (1994).
RN [4]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS PRO-478; THR-617;
RP MET-832 AND LYS-880.
RA Livingston R.J., Rieder M.J., Shaffer T., Bertucci C., Baier C.N.,
RA Rajkumar N., Willa H.T., Daniels M., Downing T.K., Stanaway I.B.,
RA Nguyen C.P., Gildersleeve H., Cassidy C.M., Johnson E.J.,
RA Swanson J.E., McFarland I., Yool B., Park C., Nickerson D.A.;
RT "NIHES-SNPs, environmental genome project, NIHES ES15478, Department
RT of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE OF 56-882.
RC TISSUE=Placenta;
RX MEDLINE=95324920; PubMed=7601454;
RA Bex G., Staes K., van Hengel J., Molemans F., Bussemakers M.J.G.,
RA van Bokhoven A., van Roy F.;
RT "Cloning and characterization of the human invasion suppressor gene E-
RT cadherin (CDH1).";
RL Genomics 26:281-289 (1995).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 172-311.
RC TISSUE=Liver;
RX MEDLINE=89031725; PubMed=3263290;
RA Mansouri A., Spurr N., Goodfellow P.N., Kemler R.;
RT "Characterization and chromosomal localization of the gene encoding
RT the human cell adhesion molecule uvomorulin.";
RL Differentiation 38:67-71 (1988).
RN [7]
RP NUCLEOTIDE SEQUENCE OF 265-392.
RC TISSUE=Liver;
RA Frixen U.H.;
RL Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.
RN [8]
RP NUCLEOTIDE SEQUENCE OF 1-16.
RX MEDLINE=94380041; PubMed=8093045;
RA Bussemakers M.J.G., Girolodi L.A., van Bokhoven A., Schaiken J.A.;
RT "Transcriptional regulation of the human E-cadherin gene in human
```



```
Db 188 VFSITGGADTPVGVFIIRERETGWLKVTEPLDRERIATVTLFSAVSSNGNAVEDPME 247
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 201 ISIIVTQNDHKPKFTQDTRGSLVGLPGTSMQVATDDEDAIYTYNGVAYSIIHQ 260
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 248 ILITVTQNDQNKPEFTQEVFKGSMEGALPGTSMVEVTATDADDDVNTYNAATYTL 307
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 261 EPKDPHLMFTIHRSTGTISVISGLDREKVPVEYTLTIQATDMGDGSTTTAVAVBILD 320
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 308 DPFLPKNMFTINRNTGVISVTTGLDRESFPYTLVVQAADLQEGSLSTATAVITVD 367
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 321 ANDNAPMFDPOKYBAHPVNAVGHVQRLVTDLDAPNSPAWRATYILIMGGDGDHFTIT 380
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 368 TNDNPPFNPPTYKGQVPEANVVIITLKVTDADAPNTPAWEAVYIIL-NDGQGFVVT 426
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 381 THPESNOGILTRKGLDFEAKNQHTLYVEVNEAPFVLKPTSTATTIVHVEDVNEAPVF 440
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 427 TNPVNDGILKTAKGLDFEAKQQYILHVAVTNVVPEVSLTSTATTIVDVLVDNEAPIF 486
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 441 VPPSKVVEVOEGIPTGPPVCVYTAEDPK-ENQKISYRIILRDPAGLAMPDSDGQVAVG 499
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 487 VPPEKRVEVSEDFVGQEIITSYTAQEDPTFMEQKITVIRWRDTANWLEINPDTGAISTRA 546
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 500 TLDREDSQVRNNIYEVWVLAMONGSPPTTGTGTLTLTLIDVNDHGPVPBPRTIICNQS 559
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 547 ELDRDEFEHVKNSTYALIIATDNGSPVATGTGTLTLIDVNDNAPIPEPRTIFFCERN 606
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 560 PVRHVLNITKDLSPHTSPQQAQITDSDIYWTAEVNE-EGDVTVLKFKLQDITDVH 618
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 607 PKQVINIADLPNTPSPETAELTHGASANWTIYNDPTQESILKPKMALEVGYKIN 666
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 619 LSLSDHGKQELTVIRATVCDCHGVETC--PGPWKGGF---ILPVLGAVLALLFLLV 672
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 667 LKLMNQNKQVTTLEVSVCDEGAAGVCRKAQVFEAGLQIPALLGLIGLILALLILL 726
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 673 LLLLVKRRKIKEPLLLPEDDTRDNVIFYEGEGGEDQDYDITQLHRGLEARPEVVLRN 732
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 727 LLLFLRRRAVVKPELLPEDDTRDNVYYDEEGGEDQDFLSQLHRLGDARPEVT-RN 785
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 733 DVATIIPTMYRPRPANDBIGNFIENLKAANTDPTAPPYDILLVFDYEGSGSDAASL 792
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 786 DVAPTLMSVPRYLPRPANDBIGNFIDENLKAADTPTAPPYDLSLLVFDYEGSGSEASL 845
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 793 SSLTSSASDQDDYDYLNEGWSRFKKLADMYGGGEDD 829
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 846 SSLSSSESDDQDDYDYLNEGWNRFKKLADMYGGGEDD 882
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 8
Q6R8F2_BOVIN PRELIMINARY; PRT; 882 AA.
ID Q6R8F2_BOVIN PRELIMINARY; PRT; 882 AA.
AC Q6R8F2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE E-cadherin.
GN Name=CDH1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hanna C.B., Cox M.L., Golding M.C., Westhuesin M.E., Kraemer D.C.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY508164; AAR01598.1; -; mRNA.
DR SRR; O6R8F2; 155-372, 802-877.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005059; F:calcium ion binding; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C term.
```

```
DR Pfam; PF00028; Cadherin; 5.
DR Pfam; PF01049; Cadherin_C; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS02268; CADHERIN_2; 4.
KW Calcium.
SQ SEQUENCE 882 AA; 97938 MW; 0EC5C82ADF7D5F36 CRC64;

Query Match 55.1%; Score 2406.5; DB 2; Length 882;
Best Local Similarity 53.8%; Pred. No. 1.3e-141;
Matches 476; Conservative 120; Mismatches 225; Indels 63; Gaps 12;

Qy 6 GPLA-SLLLLQVC-----WLQCAASBPCCRAVFREAETLEAGGAEOBPGQALGV-PMGC 58
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 2 GPWSRSLSALCCCRCPNWL-CREPEPCIPGFAESYTFVPRNLERGRVLRVSEGC 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 59 PGQEPALFSTNDNDFTVRNGETVQERRSLKERNP----- 92
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 AGLPRTVYVSDDRFKVHTDGLTVRRPVLHRLPELSFLVHAWDSTHRKLSKTVLEVSA 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 93 -----LKI FFSKRI-LRRHKRDVAVAPISVPENKGPFPQRLNOLKS 133
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 HHHHHSHHDSPGTQTEVLTFFGPHHGLRRQRQDWIPIPCPENKGPFPKSLVOIKS 180
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 134 NKDRDTKIFYSIITPGADSPPEGVFAVEKETGMLLNKPLDRDEIAKYELFGHAVSENGA 193
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 NKEKEQVFYSITQGRADTPGVFIERETGMLKYQPLDRQIAKYILFHAVSSNGQ 240
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 194 SVEDPMNISIIIVTDQNDHKPKFTQDTRGSLVGLGVIPTGTSVMQVATDEDDAIYTYNGV 253
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 AIEBPEMIVITVDQNDKNKQFTQEVFKASALEGALPGTSMQVTAIDIDEVNTYTAAI 300
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 254 AYSIHQEPKDPHDLAMFTIHRSTGTTISVSSGLDREKVPYTLTIQATDMGDGSTTAV 313
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 GYTIPAQDPLPHNKMFTINKETGVISLVTGLDRESFPTTLMVQAADLNGEGLSTAT 360
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 314 AVVEIILDANDNAPMFPQKYEAHVPENAVGHVQRLVTDLDAPNSPAWRATYILIMGGDD 373
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 361 AVITVLTNDNAPRFPNTYTVGSPENEAANVITLTVTDADDPNTFAWEAVTVL-NDN 419
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 374 GDHFTITTHPESNOGILTRKGLDFEAKNQHTLYVEVNEAPFVLKLPSTATTIVHVED 433
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 420 EKQFIVVTDVPTNEGTLTKAKGLDFEAKQQYILYVAVTNVAPFEVTLPTSTATTIVDVID 479
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 434 VNEAPVFPSPKVVVEQEGIPTGEPVCVYTAEDPK-ENQKISYRIILRDPAGLAMPDSD 492
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 480 VNEAPIFVPQKRVPEPDEFGVGLSITSYTAREPDTFMEQKITVIRWRDTANWLEINPET 539
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 493 GOVTAVGTLDRDEDEQFVRNNIYEVWVLAMONGSPPTTGTGTLTLTLIDVNDHGPVPBP 552
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 540 GAISRAELDRDVEDVHVKNSTYALIIATDNGSPATGTGTLTLFLDDVNDNGVPBPRT 599
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 553 ITICNQSPVRHVLNITKDLSPHTSPQQAQITDSDIYWTAEVN-EGDVTVLVSLKFKLK 611
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 600 MDCFQRPPEPIININPDLPNTPSPETAELTHGASVNWITIEYNDQERESLILKPKKTL 659
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 612 QDTYDVHLSSLSDHGKQELTVIRATVCDCHGVETC--PGPWKGGF---ILPVLGAVLA 665
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 660 LGDHKINLKILDNQNKQDVTTLDVHVCDGIVSNCRKARPAPAEAGLQVPAIPLGLIGLILA 719
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 666 LLFLLVLLVLLLVKRRKIKEPLLLPEDDTRDNVIFYEGEGGEDQDYDITQLHRGLEAR 725
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 720 FLILILLVLLVRRRVVVKPELPPEDDTRDNVYYDEEGGEDQDFLSQLHRLGDAR 779
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 726 PEVVLNRNDVAPTIIPTPMYRPRPANDBIGNFIENLKAANTDPTAPPYDILLVFDVEGS 785
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 780 PEVT-RNDVAPTLMSVQYRPRPANDBIGNFIDENLKAADSDPTAPPYDLSLLVFDVEGS 838
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 786 GSDAASLSSTSSASDQDDYDYLNEGWSRFKKLADMYGGGEDD 829
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 839 GSEATLSSLNSESDDQDDYDYLNEGWNRFKKLADMYGGGEDD 882
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```


[illegible]

Db	249	EIVITVTQDNDRPEFTQEVFEGSVAREGAVPGTGVSMKVKSATDADDVNTYNTAAIAYTIVS	308
Qy	260	QEPKDPDLFTTHRSVTSVSSGLDREKVPETLTICATDMGDGSGSTTTTAVAVVEIL	319
Db	309	QDPFLPHKNFTVNRDVGTVSVTSGLDRESYPTTLVUQAADQGEGLSTTAKAVITVK	368
Qy	320	DANDNAPMFDPQYEAHVPEVNAVCHVEQRIITVTDLDAPNSPAWRATVLYLIMGDDGDHFTI	379
Db	369	DINDNAPVFNPSYQGVPEVNEVNARATLKVTDGDDAPNTPAMKAVTVV-NBPDQQFVV	427
Qy	380	TTHPESNOGLITRKGLDFAKQHTLYVEVNEAPVVKLPSTATIVVHVDVNEAPV	439
Db	428	VTDPPTNDGILKTAAGLDFAKQYILHRVENEPEPESGLVSTATVTVVVDVNEAPI	487
Qy	440	FVPSKVEVQEGIPTCEPCVVTAEPPDK-ENOKISYRILRDPAGWLAMPDPSQGVTA	498
Db	488	FMPAERREVEPDPFGVQGEITSYTAREPDTFMDQKITRYIRWDTANWLEINPETGAIFTR	547
Qy	499	GTLDREDEQFVRNNIYEMVMVLMNDGSPPTTGTGTLTLLTLDVNDHGVPEPRQITICNQ	558
Db	548	AEMDREDAEHKVNSTYVALLIATDDGSPATGTGTLTLLVLDVNDNAPIPEPRNMQFCOR	607
Qy	559	SPVRHVLNITKXLSPTSFPQALTDSDIYMTAEVNEBG-DTVVLSLKFKFLKQDTYDV	617
Db	608	NPQPHIITILDPDLPNTSPTELTGASVNMWTIEYNDAAQESLILQPRKDLGEIGYKI	667
Qy	618	HLSDSHGNKEOLTIVIRATVCDCHGVETCGPMKGGF-----ILPVLGNVALLF	668
Db	668	HLKLADNQNQDQVTLTDVHVCDCGTYNNC---MKAGIVAAAGLVQVPAILGILGILALLI	724
Qy	669	LLVLLVLLVKKRRIKLEPLLPEDDTRDNYFYEGEGGEDDYDITQLHRLGLEARPEV	728
Db	725	LILLLLFLRRRTVVKELPPLPDDTRDNYFYDEEGGEDQDFDLSQLHRLGLDARPEV	784
Qy	729	VLNRDVAPTIPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVDFYEGSGSD	788
Db	785	T-RNDVAPTLMSVQYRPRPANPDEIGNFIDENLKAADSDPTAPPYDLSLVDFYEGSGSE	843
Qy	789	AASLSSLTSSASDQDYDYLNEGSRFKLADMYGGEDD	829
Db	844	AASLSSLSNESDQDQDYDYLNEGSRFKLADMYGGEDD	884
RESULT 12			
Id	CADH1	RAT	
AC	CADH1	RAT	STANDARD; PRT; 886 AA.
DT	Q9ROT4		O35794; Q9JIV9;
DT	16-OCT-2001		(Rel. 40, Created)
DT	16-OCT-2001		(Rel. 40, Last sequence update)
DT	13-SEP-2005		(Rel. 48, Last annotation update)
DE	Epithelial-cadherin precursor (E-cadherin) (Uvomorulin) (Cadherin-1).		
CN	Name=Cdhl;		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
OC	Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
FN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=Sprague-Dawley; TISSUE=Jejunum;		
RA	Asai K., Tada T., Yamamoto M., Obayashi M., Mizuno M., Toda A.,		
RA	Einoto T., Kato T.;		
RL	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.		
FN	[2]		
RP	NUCLEOTIDE SEQUENCE OF 769-872.		
RC	STRAIN=Fischer 344; TISSUE=Testis;		
RX	MEDLINE=20114371; PubMed=10650949; DOI=10.1210/en.141.2.675;		
RA	Johnson K.J., Patel S.R., Boekelheide K.;		
RT	"Multiple cadherin superfamily members with unique expression profiles		
RT	are produced in rat testis.";		
RL	Endocrinology 141:675-683(2000).		
FN	[3]		
RP	NUCLEOTIDE SEQUENCE OF 779-846.		
RC	STRAIN=Mammary tumor;		
RA	Gibbons K.L.;		
RA	Thesis (1997), University of Technology / Sydney, Australia.		
CC	-I- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.		
CC	They preferentially interact with themselves in a homophilic		
CC	manner in connecting cells; cadherins may thus contribute to the		
CC	sorting of heterogeneous cell types. E-cadherin has a potent		
CC	invasive suppressor role.		
CC	-I- SUBUNIT: Homodimer. Binds CTNNB1 (By similarity).		
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).		
CC	-I- SIMILARITY: Contains 5 cadherin domains.		
CC	-----		
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use as long as its content is in no way modified and this statement is not		
CC	removed.		
CC	-----		
DR	EMBL; AB017696; BAA84920.1; -; mRNA.		
DR	EMBL; AF177680; AAF87055.1; -; mRNA.		
DR	EMBL; AJ000540; CAA04173.1; -; mRNA.		
DR	HSP; P09803; I17X.		
DR	SMR; Q9ROT4; 159-376, 806-881.		
DR	Ensembl; ENSRNOG0000020151; Rattus norvegicus.		
DR	GO; GO:0007416; P:synaptogenesis; IMP.		
DR	InterPro; IPR002126; Cadherin.		
DR	InterPro; IPR000233; Cadherin_C_term.		
DR	Pfam; PF00028; Cadherin; 5.		
DR	Pfam; PF01049; Cadherin_C; 1.		
DR	PRINTS; PR00205; CADHERIN.		
DR	SMART; SM00112; CA; 4.		
DR	PROSITE; PS00232; CADHERIN_1; 3.		
DR	PROSITE; PS0268; CADHERIN_2; 5.		
KW	Calcium; Cell adhesion; Glycoprotein; Phosphorylation; Repeat; Signal;		
KW	Transmembrane.		
FT	SIGNAL	1	23 Potential.
FT	PROPEP	24	158 Potential.
FT	CHAIN	159	886 Epithelial-cadherin.
FT	TOPO_DOM	24	713 Extracellular (Potential).
FT	TRANSHEM	714	734 Potential
FT	TOPO_DOM	735	886 Cytoplasmic (Potential).
FT	DOMAIN	159	266 Cadherin 1.
FT	DOMAIN	267	379 Cadherin 2.
FT	DOMAIN	380	490 Cadherin 3.
FT	DOMAIN	491	597 Cadherin 4.
FT	DOMAIN	598	701 Cadherin 5.
FT	COMBIAS	842	857 Ser-rich.
FT	MOD_RES	842	842 Phosphoserine (By similarity).
FT	MOD_RES	844	844 Phosphoserine (By similarity).
FT	MOD_RES	850	850 Phosphoserine (By similarity).
FT	CARBOHYD	562	562 N-linked (GlcNAc...) (Potential).
FT	CARBOHYD	641	641 N-linked (GlcNAc...) (Potential).
SQ	SEQUENCE	886 AA;	98715 MW; A9AE28EB797A547 CRC64;
Query Match			
Best Local Similarity 54.7%; Score 2388; DB 1; Length 886;			
Matches 477; Conservative 122; Mismatches 212; Indels 76; Gaps 14;			
Qy	9	ASLLLLQV-CWLQC-----AASEPCRAVFRAEAVTELAGGAEOEPQALGKV-FMGCPQOE	62
Db	10	ALLLLQVSSWL-CQPESESDSCRPGFSSEVYTFVLPERHLERGHILGRVKFEGCTGRP	68
Qy	63	PALFSTNDPFTVRNGETVQERSLK-----	88
Db	69	RTAFSEDSRPFKVGSTDGVITVKRHLKHLKLETSLFVANDSSYRKLSTKVTLSLGHVHH	128
Qy	89	-----ERNP-LKIFPS-KRILRRKRDVWVAPISVPENKGPFPQRLNQKSNKDR	137
Db	129	RHHHRDPVSESNPELLTFPSFHQGLRQKRDWVIPPINCENQKGEFPQRLVQIKSNRDK	188
Qy	138	DTKLFYSITGPGADSPGEGVFAVEKETCWLILNKPLDREETAKYELFGHVAENGASVED	197
Db	189	ETTVFYSITGPGADKPPGVGVFIIBRETGWLWKTQPLDREADIKYLLYSHAVSNGEAVED	249

QY 198 PMNISIIVTDNDHKKFTQDTRGSLVGLPCTSMQVATATDEDDAIYTYNGVAVYSI 257
Db 249 PMEIVVTVDQNDNRBEFIOEVPEGSVAEGALPCTSMQVSATDADDDINTYNAAYTI 308
QY 258 HSOEPKDPHDMETIHRSTGTSIVISGLDREKVPVETLTIOATDMGDSITTAIVAVE 317
Db 309 LSOQPELPHKMTFVARDTGIVSVTSGLDRESYPTTLLVVQADIQEGELSTTAKAVIT 368
QY 318 ILDANDNAPMFDQKVEAHVPENAVGHEVORLTVTDLDPNSPAWRATYLMGDDGDH- 376
Db 369 VKDINDNAPINFSTVQGVLENEVGARIATLKVTDDDAPNTAWNAVTVV--NPDHQ 426
QY 377 FTITTHPESNQGLITTRKGLDFAKQOHTLYVEVNEAPVCLKPTSTATIVVHVVDNE 436
Db 427 FTVITDPKTNIGLTKAKGLDFAKQOYILHVTVENEPEPESLVPSTATVTVDDVVDNE 486
QY 437 APVFVPKSKVEVOEGIPTEGPEVCVVTAEEDPK-ENOKISYRILRDPAGHLANDPDSGV 495
Db 487 APIFVPAEKRVVEPDEFGVLEIASYTAREPDTFMEQKITRIRWDTANWLEINPETGVI 546
QY 496 TAVGTLDREDEQFVRNNIYEVMLAMDNQSPPTTGTGLTLLTLDVNDHGPVPEPRQITI 555
Db 547 STRAENDREDSERVKNSTYTALIIATDDGSPATGTGTLTLLVSDVNDNAPIPEPRNQF 606
QY 556 CNOQSPVRHVUNITKXLSHPTSPFQALTDSDSIYMTAEVNE-EGDTVVLKXKFLKQDT 614
Db 607 CORNPXPHVITLDPOLPNTSPFTAEHTGASVNNWTIEYNDAEQSLILQPKKDLKIGS 666
QY 615 YDVHLSLSHGNKEOLTIVIRATVCDCHGHVETC-----PGPKGGPILPVGA 662
Db 667 YKINKLSNQNQKQDTTLEHVHVCDEGTVNNCMKAISLEAGLQVPA-----ILGILGG 720
QY 663 VLALLFLVLLVLLVRKKRIKEPLLLPEDDTRDNNVYFYEEGGEGGEDODYDITQLHRGL 722
Db 721 ILALLLILLLFLRRRTVWKEPLFPDDTRDNNVYFYEEGGEGGEDODFSLQSLHRGL 780
QY 723 EAPPEVLNDVAPTIIPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDITLAVFY 782
Db 781 DARPEVI-RNDVAPTLMSVPQYRPRPANPDEIGNFIDENLKAADSPTAPPYDLSLAVFY 839
QY 783 EGSQSDAASLSSTSSASODODYDYLNEGSRFKKLADMYGGGEDD 829
Db 840 EGSQSEAASSLSSNESDQDQDYDYLNEGSRFKKLADMYGGGEDD 886

RESULT 13
Q5RAX1_PONPY
ID Q5RAX1_PONPY PRELIMINARY; PRT; 882 AA.
AC Q5RAX1;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFp469A211.
GN Name=DKFp469A211;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RG The German cDNA Consortium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Oeinger A., Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR858890; CAH91089.1; ; mRNA.
DR SWR; Q5RAX1; 155-371.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR InterPro; IPR002126; Cadherin.

DR InterPro; IPR002126; Cadherin_C_term.
DR Pfam; PF00028; Cadherin; 5.
DR Pfam; PF01049; Cadherin C; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00322; CADHERIN_1; 3.
DR PROSITE; PS0268; CADHERIN_2; 4.
KW Calcium; Hypothetical protein.
SQ SEQUENCE 882 AA; 97427 MW; 3EC35657C1B6C719 CRC64;

Query Match 54.6%; Score 2387; DB 2; Length 882;
Best Local Similarity 54.5%; Pred. No. 2.2e-140; Indels 60; Gaps 13;
Matches 478; Conservative 114; Mismatches 225;

QY 9 ASILLIQV-CWLCQAASEPCRAVFREAETVLEAGAEQEPGQALGV-FWGCPG-QEPAL 65
Db 10 ALLLLLQVSSWL-CQEPCHGCFDAESYFTVPRHLEGRVILGRVFNEDCTGRQRTAY 68
QY 66 FSTDNDDDFTVRNGETVQERRSLKERNP----- 92
Db 69 FSLDT-RFKVGTGDTVITVKRPLRFHNPQIHFLVYAWDSTYRKFTKVTLTNTVGHHRPLP 127
QY 93 -----LKIIP-SKRILRRHKRDWVAPISVPENKGGPFPQRLNQLKSKDRDTK 140
Db 128 HQASVSGIQAEULTFPNSSGSLRRRRKRDWVIPPISCPENEKGGPFPKLVQIKSKDKGK 187
QY 141 IFYSITGPGADSPGCVFAVEKETGMLLNLKPLDREIYAKYELFGHAVSENGASVEDPMN 200
Db 188 VFYSITGQADTPVGVFIIRERTGWLKVTPELDRERIATYTLFHAVSSNGNAVEDPME 247
QY 201 ISIIVTQNDHHPKFTQDTPFRGSLVGLVPGTSMQVATATDEDDAIYTYNGVAVYSIHQ 260
Db 248 ILITVTDQNDKPEFTQEVFKGSMEGALPGTSMVEVTATDADDDVNTYNAAYTILSQ 307
QY 261 EPKDHLMFTTHRSNTGISVSSGLDREKVPVETLTIOATDMGDSITTAIVAVEILD 320
Db 308 DPELDPKXNFTINRNTGVISVVTGLDRESFTYTLVQAAQLQSGELSTTATVITVD 367
QY 321 ANDNAPMFDQKVEAHVPENAVGHEVORLTVTDLDPNSPAWRATYLMGDDGDHFTIT 380
Db 368 TNDNPPVNPITYKQVPEDEANVTTLKVTDADAPSTPAWEAVYTL-NDNGQGVVT 426
QY 381 THPESNQGLITTRKGLDFAKQOHTLYVEVNEAPVCLKPTSTATIVVHVVDNEAPVF 440
Db 427 TNPVNDGILTKAKGLDFAKQOYILHVAVTNVVPFVSLTSTATVTVDDVLDVNEAPIF 486
QY 441 VPSPKVEVOEGIPTEGPEVCVVTAEEDPK-ENOKISYRILRDPAGHLANDPDSGVTA 499
Db 487 VPPEKRVESDFGVGQEITSYTAWEPTFMEQKITRIRWDTANWLEINPDTGAISTRA 546
QY 500 TLDREDEQFVRNNIYEVMLAMDNQSPPTTGTGLTLLTLDVNDHGPVPEPRQITCNOS 559
Db 547 ELDRDREHVKNSTYTALIIATDNGSPVATGTGTLTLLVSDVNDNAPIPEPRTLFCERN 606
QY 560 PVRHVNLITDKLSHPTSPFQALTDSDSIYMTAEVNE-EGDTVVLKXKFLKQDTYDVH 618
Db 607 PKPQVINIADLPNTSPFTAEHTGASVNNWTIQVNDPTQSSIILKPKMALEVGVDYKIN 666
QY 619 LSLSDHGNKEOLTIVIRATVCDCHGHVETC--PGPKGGF-----ILPVGLAVALLPLL 672
Db 667 LKMDNQNQKQDTTLEVGVCDCGCVAGVCKKQAPTEAGLQIPAILGILGILLALLLIL 726
QY 673 LLLLVKRRKRIKEPLLLPEDDTRDNNVYFYEEGGEGGEDODYDITQLHRGLARPVEVL 732
Db 727 LLLFLRRVAVKEPLLPEDDTRDNNVYFYEEGGEGGEDODFSLQSLHRGLDARPEVT-RN 785
QY 733 DVAPTIPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDITLAVFYEGSGSDAASL 792
Db 786 DVAPTLMSVPYLRPRPANPVEIGNFIDENLKAADTDPTAPPYDLSLAVFYEGSGSEASL 845
QY 793 SSLTSSASODODYDYLNEGSRFKKLADMYGGGEDD 829
Db 846 SSLNSSESDKQDQDYDYLNEGSRFKKLADMYGGGEDD 882


```
RESULT 14
Q6NTM0_XENLA
ID Q6NTM0_XENLA PRELIMINARY; PRT; 884 AA.
AC Q6NTM0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Atschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Stapleton M., Marusina K., Rubin A.A., Casavant T.L., Hong L.,
RA Datchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Tohivuyki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391 (2002).
[3]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Embryo;
RC EMBL; BC068940; AAH68940.1; -, mRNA.
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391 (2002).
[3]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Embryo;
RC PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS00268; CADHERIN_2; 4.
KW Calcium, Hypothetical protein.
SQ SEQUENCE 884 AA; 98014 MW; 5321B1B5E2FC635F CRC64;
```

Query Match 54.4%; Score 2377; DB 2; Length 884;
Best Local Similarity 54.2%; Pred. No. 9.2e-140;
Matches 470; Conservative 125; Mismatches 221; Indels 51; Gaps 8;

```
QY 11 LLLQVCLQCAASEPCRAVFREAEVTLBAGGAEPGQALGV-FMGCPGQBPALFSTD 69
DB 18 LCLQVVPISINVDVSCQPGFSSANVTFSVNRLELGRKLGKGVNFVDCITRKHGLVDVG 77
QY 70 NDDFTVRNGETV-----QERSLK 88
DB 78 DSRFRVLPDGTVLVLRHVKLHSLKSDTRFTISTWDARGIKHSTNISVVNKRHSRSEARSRS 137
QY 89 ERNPLKIFPSKRI-LRRHKRDVWVAPISVPENGKGPFPORLNOLKNKORDTKIFYSITG 147
DB 138 SELPVITTFPEKHTGLRKRRDWIPKIKVSENERGPPFKRLVQIKNKKEKLSKVFISITG 197
QY 148 PGADSPPEGVFAVEKETGMLLNKPLDRBEIAKYELFEGHVASENGASVEDPMNISIVTD 207
DB 198 QGADTPPEGIFRIEKEKTMQVTRPLDREYEKVVLLSHAVSENGASVEEPMETVITVID 257
QY 208 QNDHKPKFTQDTPRGVLEGLVPGTSVMQVATDEDDAIYTYNGVAVYSIHSEPKDPHD 267
DB 258 QNDNRPKFTQPVFRGVSREGVQGTVMVSATDDDDSIDSLNGVIAYSILKQDPEPIP 317
QY 268 LMFTIHRSTCTISVISGLDREKVPETLTIOATMDGSGTTTAVAVVEILDANDNAPM 327
DB 318 NLFETINRETCVISLIGTLDRKFPETTLTVQAADLDGAGLTAEKAKAVISITDANDNAPI 377
QY 328 FDPQKYEAHVPENAVGEVQRLTVTDLDAFNSPAMRATYLLIMGDDGDGDFHTIITHPESNQ 387
DB 378 FDPKTYTALVPENEVGEVQRLSVTDLDMFGTAQAQAVYKIR-VNEGFFENITDPESNQ 436
QY 388 GIITTRKGLDFAKNOHTLYVEVTNEAPFVLKLPSTATIVVHVEDVNEAPVFPVSKV 447
DB 437 GILTTAKGLDFEVKQYVIQITVENAVPFSVPLPTSTATVTVTVEDVNEAPVFPVVSRY 496
QY 448 EVOEGITPCEPVCVYTAEPDKEN-OKISVRIILRDPAGWLAMPDSDGQVTAAGTLDREDE 506
DB 497 DVSEDLTRGEKIVSLVAQDPDKQIQKLSYFIGNDPARMLTINKDNGIIVTGNGLDRESE 556
QY 507 QFYRNNIYEVWLAMDNGSPPTTGTGTLTLTLVDNDHGFVPFPPTQITICNQSPVRHVLN 566
DB 557 -YVKNTYTVIMLVTDGVPVGTGTGLILHLVDINDNGFPVSPRFTMCDQNPQVLT 615
QY 567 ITDKLSPTSPPQAOITDDSDIYWTAEVNEEGDVTVLKFLKQDYDVHLSLSHGN 626
DB 616 ITDADIPNTYPSVSLSHGSELTAKAELDSKGTSMRLSPTOQLKGDYGIYVLLADAQA 675
QY 627 KEOLTIVIRATVCDCHGVETCPGPKGGFILP----VLGAVLALLFLLVLLVLRKRX 682
DB 676 NRQLTVNATVCICEGRAIKCQKLVAGFDLPILVILGSILALLLILSLLLFLKRRKV 735
QY 683 IKEPLLLPEDDTRDNVFFYEGEGGEDDQDYITQLHRGLEARPEVVLNRNDVAPTIIPTP 742
DB 736 VKEPLLLPEDDTRDNVFFYEGEGGEDDQDYDLSQLHRGLDARPD-IMRNDVVPITMSVP 794
QY 743 MYRPRPANPEIGNFIENLKAAWTDPTAPPYDTLLVDFYEGSGSDAASISLSTSSASDQ 802
DB 795 HYRPRSPNPEIGNFIDENLDAADNDPTAPPYDLSLVDFYEGSGSEAAASLSSNSNN 854
QY 803 DQDYDYLNEGSRFKKLADMYGGEDD 829
DB 855 EHDYNYLNDWGPFRKLADMYGGDDDD 881
RESULT 15
CADHP_XENLA
ID CADHP_XENLA STANDARD; PRT; 880 AA.
AC P33148; Q91543;
DT 01-OCT-1993 (Rel. 27, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Ep-cadherin precursor (C-cadherin).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
```

S -> P (in Ref. 2).
S -> P (in Ref. 2).

573	573
864	864
165	162
178	174
183	183
185	189
196	189
199	197
200	200
204	203
208	206
210	210
212	211
217	214
223	223
225	223
237	228
242	242
247	247
254	254
262	262
263	268
273	274
275	274
284	280
288	287
299	298
308	302
320	318
320	318
324	322
329	326
338	336
343	341
353	350
354	354
354	354
359	357
367	362
376	376
380	380
386	393
397	393
399	399
406	405
409	408
412	412
413	414
416	416
420	420
421	420
428	424
432	430
439	435
455	451
458	458
472	472
477	486
487	482
492	491
496	496
499	498
504	504
511	511
512	516
517	516
523	522
536	535
537	538
541	542
545	542
552	551
562	562
569	579
582	579
585	585
595	595
602	600
614	610
614	610
620	620
629	629
639	639
642	642
645	644
652	648
662	662
665	665
667	667

FT	STRAND	676	679	
FT	STRAND	682	684	
SQ	SEQUENCE	880 AA; 97651 MW; C6CCD91566427D86 CRC64;		
Query Match 54.4%; Score 2375; DB 1; Length 880;				
Best Local Similarity 54.2%; Pred. No. 1.2e-139;				
Matches 471; Conservative 128; Mismatches 214; Indels 56; Gaps 9;				
Qy	11	LLLLQVCLQCAASEPRAVEATLEAGAEQEQALGV-FWGCGPQEPALFSTD	69	
Db	17	LCLLQVPSINADVSGCKPGFSSAEYIFSVNRRELERKRLGVNFSDCTTRKHGLDVG	76	
Qy	70	NDDFTV-----RNGETVQERRS	86	
Db	77	DSRFRLPDGTVLVKRVHKLHKDTFTISTWDARGIKHSTNIAVASKRHSRSGEEAHSRS	136	
Qy	87	LKERNPLKIPPSKRI-LRRHKROWVAPIVSPENGKGPFFORLNQLKSNKORDTKIFYSI	145	
Db	137	---KLPVLTFPETHTGLKRRKRDWVIPPVKYSENERGPPFKRLVQIKSNKDRFNKVYYSI	193	
Qy	146	TGPCADSPPEGVFAVEKETGWLANKPLDREEIAKVELFGHAVSENGASVEDPMNISIIV	205	
Db	194	TQQGADNPPQGVFRIEWETGMLVTRPLDREEDKYVLSSHAVSENGSPVEEPMETINV	253	
Qy	206	TDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVTTATDEDDAIYTYNGVVAYSIHSQEPKOP	265	
Db	254	IDQNDNRPKFTQDVFRGSVREGVQGTQVMASATDEDDNIDSLNGVLSYILKQDPEEP	313	
Qy	266	HDLMTTHIRSTGTTISSGLDREKVPYTLTIQATDMOGDGSTTTAAVAVVAILDANDNA	325	
Db	314	IPNLFTINRETGVISLIGTGLDREKFPYTLTVQATDLEGAGLSVEGKAIQITDANDNA	373	
Qy	326	PMFDQPKYEAHVPENAVCHVORLTVTDLDAPNSPANRATVYLMGGDDGDHFTITTHPES	385	
Db	374	PIFDPKTYTALVPENEIGFEVQRUSVTDLDMPGTPAWQAVYKIR-VNEGGFNITTDPE	432	
Qy	386	NOGILTTKGLDFAKQOHTLYEVETNEAPPVLKLPSTATIVVHVVEDVNEAPVFPVPSK	445	
Db	433	NOGILTTAKGLDFELRKQVLOITVENAEPSPVPLPTSTATVTVTVEDVNEAPFFVPAVS	492	
Qy	446	VVEVOEIGPTGEPVCVYTAEDPKEN-QKISYRILRDPAGWLAMPDSGGQVTAVGTLDRE	504	
Db	493	RVDVSEDLRGEKIISLVAQDPDKQIQKLSYFCNDPARWLTVNKONGIVTGNGLDRE	552	
Qy	505	DEQFVRNIVEMVLAMDNQSPPTGTGCTLLTLIDVNDHGPVPEPRQITICNOSPVRHV	564	
Db	553	SE-YVKNTTYTIVMLVTDGVSVGTGTGLTLHLVDVNDNGVPSPRVFTMCDQNPEPQV	611	
Qy	565	LNITDKLSPTSPPQAOLTDSDIYMTAEVNEEGDVTVLSLKKFLKQDTYDVHLSLSDH	624	
Db	612	LTISDADIPNTYKYVLSHGSDLTWKAELDSKTSMLLSPTQQLKKGDYSIYVLLSDA	671	
Qy	625	GNKEQLTVIRATVCDCHGVETCPGPKWGGFILP-----VLGAVLALLFLLVLLLVKK	680	
Db	672	QNNPQLTVVNATVCSCEGAIKCQEKLVGGFDLPILVILGSVLALLILFLLLLFLKRK	731	
Qy	681	RKIKEPLLPEDDTRDNVYVYEGEGGEEDDYDITOLHRLGLARPEVVLRLNDVAPTIIP	740	
Db	732	KVWKEPLLPEDDTRDNIIFYYGERGGEEDDYDLSQLHRLGDSRPD-IMRNDVVPITMP	790	
Qy	741	TPMYRPRPANDPDEIGNFTIENLKAANTDPTAPPYDTLLVFDYEGSGSDAASLSLTSSAS	800	
Db	791	APHYRPRPSNDEIGNFDENLDAADNDPTAPPYDSLLVFDYEGSGSEASLSLSNSNS	850	
Qy	801	DODQDYDYNLWNGSRFKKLADMYGGEDD	829	
Db	851	NDEHDYNYLSDWGSFRKRLADMYGGDDDE	879	


```
Db 1 MELLSPHAFULLLLQVCLRSVVSEPYRAGFIGEAGVTLEVEGTDLSPQVLGKVALAG 60
Qy 59 PGQEPALFSTNDNDFTVRNGETVQERSLKERPLKIFPSKRILRRHKRWVAVPISVPE 118
Db 61 QGMHHA----DNGDIIMLTGTVGGKDAMHS-----PPTRIILRRKRWWMPPIFVPE 110
Qy 119 NGKGPFFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPDLREEI 178
Db 111 NGKGPFFQRLNQLKSNKDRGTKIFYSITGPGADSPPEGVFTIEKESGWLILHWPDLREKI 170
Qy 179 AKYELFHAVSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGSVLGVLPGTSVMQVT 238
Db 171 VKYELGHAVSENGASVEEPMNISIIITDQNDKPKFTQDTFRGSVIEGVMPGTSVMQVT 230
Qy 239 ATDEDDAIYNGVAVSIHQSQBPKDHLMTFTHRSTGTISVSSGLDREKVEYTLTI 298
Db 231 ATDEDDAVTYNGVAVSIHQSQBPKDHLMTFTHRSTGTISVSSGLDREKVEYRLTV 290
Qy 299 QATDMDCGSGTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHEVQRLTVTDLDPN 358
Db 291 QATDMDCGSGTTTAAVAVVQILDANDNAPEPEPOKYEAVPENEVEGHEVQRLTVTDLDPN 350
Qy 359 SPAWRATYLVHVEDVNEAPVFPVPPSKVVEQEGIPTGEPVCVVTAEADPKENOKISYRI 418
Db 351 WPAMRATYHVHVEDVNEAPVFPVPPSKVVEQEGIPTGEPVCVVTAEADPKENOKISYRI 410
Qy 419 KLPTSTATIVVHVEDVNEAPVFPVPPSKVVEQEGIPTGEPVCVVTAEADPKENOKISYRI 478
Db 411 KLPTATATVVHVHVEDVNEAPVFPVPPSKVVEQEGIPTGEPVCVVTAEADPKENOKISYTI 470
Qy 479 LRDPAGLAMPDPSGQVTAAGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLL 538
Db 471 SRDPANLAMPDPSGQITAAAGILDRDEDEQFVKNNVYEVWVLATDSGNPPTTGTGTLTLL 530
Qy 539 IDVNDHGFVPEPRQITTCNOSPVRHVLNITDKOLSPHTSPFQAQLTDDSDIYWTAEVNEE 598
Db 531 TDINDHGFPEPRQIICNOSPVPQVLNITDKOLSPNSPPFQAQLTHDSDIYMAEVSEK 590
Qy 599 GDTVVALSKKFLKODTYDVLHLSLDHGNREQLTMRATVCDCHGV-ETCPGPKGGFIL 657
Db 591 GDTVALSKKFLKODTYDVLHLSLDHGNREQLTMRATVCDCHGV-ETCPGPKGGFIL 650
Qy 658 PVLGAVLALLFLLVLLVLRKKRKEIPELILPDDTRDNVYFYGEGGGEEDQDYITQ 717
Db 651 PILGAVLALLTLLALLVLRKKRKEIPELILPDDTRDNVYFYGEGGGEEDQDYITQ 710
Qy 718 LHRGLEARPEVLRNDVAPTIPTPMYRPRPANDEIGNFIENLKAANTDPTAPPYDTL 777
Db 711 LHRGLEARPEVLRNDVVPTEPTPMYRPRPANDEIGNFIENLKAANTDPTAPPYDSL 770
Qy 778 LVFDYEGSGDAASLSLTSASDQDQDYDYLNEWGSFRFKLADMYGGGDD 829
Db 771 MVFDYEGSGDAASLSLTSASDQDQDYDYLNEWGSFRFKLADMYGGGDD 822
```

RESULT 2

```
US-08-474-068A-9
; Sequence 9, Application US/08474068A
; Patent No. 5837525
; GENERAL INFORMATION:
; APPLICANT: Ranscht, Barbara
; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,068A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/213,361
; FILING DATE: 14-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/607,293
; FILING DATE: 30-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1683
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-474-068A-9

Query Match 81.3%; Score 3553.5; DB 1; Length 822;
Best Local Similarity 81.5%; Pred. No. 0;
Matches 678; Conservative 58; Mismatches 83; Indels 13; Gaps 5;

Qy 1 MGLPRGPLA-SLLLLQVCLQCAASEPCRAVF-REAEVTLAEAGAEQEQALGKVPFGC 58
Db 1 MELLSPHAFULLLLQVCLRSVVSEPYRAGFIGEAGVTLEVEGTDLSPQVLGKVALAG 60
Qy 59 PGQEPALFSTNDNDFTVRNGETVQERSLKERPLKIFPSKRILRRHKRWVAVPISVPE 118
Db 61 QGMHHA----DNGDIIMLTGTVGGKDAMHS-----PPTRIILRRKRWWMPPIFVPE 110
Qy 119 NGKGPFFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPDLREEI 178
Db 111 NGKGPFFQRLNQLKSNKDRGTKIFYSITGPGADSPPEGVFTIEKESGWLILHWPDLREKI 170
Qy 179 AKYELFHAVSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGSVLGVLPGTSVMQVT 238
Db 171 VKYELGHAVSENGASVEEPMNISIIITDQNDKPKFTQDTFRGSVIEGVMPGTSVMQVT 230
Qy 239 ATDEDDAIYNGVAVSIHQSQBPKDHLMTFTHRSTGTISVSSGLDREKVEYTLTI 298
Db 231 ATDEDDAVTYNGVAVSIHQSQBPKDHLMTFTHRSTGTISVSSGLDREKVEYRLTV 290
Qy 299 QATDMDCGSGTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHEVQRLTVTDLDPN 358
Db 291 QATDMDCGSGTTTAAVAVVQILDANDNAPEPEPOKYEAVPENEVEGHEVQRLTVTDLDPN 350
Qy 359 SPAWRATYLVHVEDVNEAPVFPVPPSKVVEQEGIPTGEPVCVVTAEADPKENOKISYRI 418
Db 351 WPAMRATYHVHVEDVNEAPVFPVPPSKVVEQEGIPTGEPVCVVTAEADPKENOKISYRI 478
Qy 419 KLPTSTATIVVHVEDVNEAPVFPVPPSKVVEQEGIPTGEPVCVVTAEADPKENOKISYRI 478
Db 411 KLPTATATVVHVHVEDVNEAPVFPVPPSKVVEQEGIPTGEPVCVVTAEADPKENOKISYTI 470
Qy 479 LRDPAGLAMPDPSGQVTAAGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLL 538
Db 471 SRDPANLAMPDPSGQITAAAGILDRDEDEQFVKNNVYEVWVLATDSGNPPTTGTGTLTLL 530
Qy 539 IDVNDHGFVPEPRQITTCNOSPVRHVLNITDKOLSPHTSPFQAQLTDDSDIYWTAEVNEE 598
Db 531 TDINDHGFPEPRQIICNOSPVPQVLNITDKOLSPNSPPFQAQLTHDSDIYMAEVSEK 590
Qy 599 GDTVVALSKKFLKODTYDVLHLSLDHGNREQLTMRATVCDCHGV-ETCPGPKGGFIL 657
Db 591 GDTVALSKKFLKODTYDVLHLSLDHGNREQLTMRATVCDCHGV-ETCPGPKGGFIL 650
```

```
QY 658 PVLGAVLALLFLVLLVLLVLRKKRKKEPILLPDDTRDNVFFYBEGGGEEDQDYDTQ 717
Db 651 PILGAVLALLTLLALLVLRKKRKKEPILLPDDTRDNVFFYBEGGGEEDQDYDTQ 710
QY 718 LHRGLEARPEVLRNDVAPITIIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDIL 777
Db 711 LHRGLEARPEVLRNDVAPITIIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDSL 770
QY 778 LVFDYEGSGDAASLSLTSASDQDDYDYLNEWGSRPKKLADMYGGGEDD 829
Db 771 MVFDYEGSGDAASLSLTSASDQDDYDYLNEWGSRPKKLADMYGGGEDD 822

RESULT 3
US-08-472-481-8
; Sequence 8, Application US/08472481
; Patent No. 5863804
; GENERAL INFORMATION:
; APPLICANT: Ranscht, Barbara
; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,481
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/213,361
; FILING DATE: 14-MAY-1994
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1686
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-472-481-8

Query Match 81.3%; Score 3553.5; DB 1; Length 822;
Best Local Similarity 81.5%; Pred. No. 0;
Matches 678; Conservative 58; Mismatches 83; Indels 13; Gaps 5;

QY 1 MGLPRGFLA-SLLLLQVLCQAASEPCRAVVF-REAEVTLAAGGAEQPGQALGKVFMG 58
Db 1 MELLSGPHAFLLLLQVLCWLRSSVSEPYRAGFTGEAGVTLVEGTDLEPQSVLQKVALAG 60
QY 59 PGQEPALFSTDNDQDFTVRNGETVRRKSLKERNPLKIFPSKILRRHRKRDVWVAPISVPE 118
Db 61 QGMHHA---DNGDIIMLTGRGTQGGKAMHS-----PPTRILRRKRREWVMPPIFVPE 110
QY 119 NGKGPPQRNLQKSNKDRQTKIFYSITGPGADSPPEGVFAVEKFTGWLILLNPLDREBI 178
Db 111 NGKGPPQRNLQKSNKDRQTKIFYSITGPGADSPPEGVFTIEKESGWLILLHMLDREKI 170
```

```
QY 179 AKYELFCHAVSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGSLVLEGLVPGTSMQVTT 238
Db 171 VKYELVGHAVSENGASVEPMNISIIITDQNDNPKFTQDTFRGSLVIEGVMPGTSMQVTT 230
QY 239 ATDEDDAIYTVNGVVAISHSQEPKPDHLMFTIHRSTGTISVISGLDREKVPETLTI 298
Db 231 ATDEDDAVNTYTVNGVVAISHSQEPKPDHLMFTIHKSTGTISVISGLDREKVPETLTI 290
QY 299 QATDMGDGSTTTAVAVVEILDANDNAPMFDQPKYEAHVPENAVGHEVORLTVTDLDAFN 358
Db 291 QATDMGDGSTTTAEAVVOILDANDNAPFEQKYEAWPENAVGHEVORLTVTDLDPVN 350
QY 359 SPARATYLLIMGGDDGDHFTIITHPESNQILITTRKGLDPEAKNQHTLVYVETNEAPFVL 418
Db 351 WPAWRATYHIVGGDDGDHFTIITHPETNQVLTTKKGLDPEAQDQHTLVYVETNEAPFAV 410
QY 419 KLPTSTATIVHVEDVNEAPVFPSPKVEVEGEGIPTGEPVCVYTAEDPKENOKISYRI 478
Db 411 KLPTATATVVHVHVDNEAPVFPSPKVEAQEGISIGELVCITYTAQDPKEDQKISYTI 470
QY 479 LRDPAGWLAMPDSGOVTAVGTLDREDEQFVRNNIYVNMVLAMDNQSPPTTGTTLTLL 538
Db 471 SRDPANWLAVDPSGOITRAGILDRDEQFVKNNYVNMVLATDSGNPPTTGTTLTLL 530
QY 539 IDVNDHGPVPEPRQITICNQSPVRHVLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEE 598
Db 531 TDINDHGPVPEPRQITICNQSPVQVLNITDKDLSNPSFPFOAQLTHDSDIYMAEVSSEK 590
QY 599 GDTVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGV-ETCPGPKGGFLL 657
Db 591 GDTVALSLKKFLKQDTYDLHLSLSDHGNREQLTMRATVCDCHGVFNDCPRPKGGFLL 650
QY 658 PVLGAVLALLFLVLLVLRKKRKKEPILLPDDTRDNVFFYBEGGGEEDQDYDTQ 717
Db 651 PILGAVLALLTLLALLVLRKKRKKEPILLPDDTRDNVFFYBEGGGEEDQDYDTQ 710
QY 718 LHRGLEARPEVLRNDVAPITIIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDIL 777
Db 711 LHRGLEARPEVLRNDVAPITIIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDSL 770
QY 778 LVFDYEGSGDAASLSLTSASDQDDYDYLNEWGSRPKKLADMYGGGEDD 829
Db 771 MVFDYEGSGDAASLSLTSASDQDDYDYLNEWGSRPKKLADMYGGGEDD 822

RESULT 4
US-08-474-067-8
; Sequence 8, Application US/08474067
; Patent No. 5811518
; GENERAL INFORMATION:
; APPLICANT: Ranscht, Barbara
; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,067
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/213,361
; FILING DATE: 14-MAY-1994
```



```
QY 200 NISIIVTQNDHKKPKTQDTFRGSLVLEGLVPGTSVMQVTTATDEDDDAIYTYNGVAVYSIHS 259
Db 249 EIVITVTDQNDNRPEFTQEVFEGSVAGAVPGTSVMKVSATADDDVNTYNAAIYTVS 308
QY 260 QEPKDPHDLMTFTHRTGTISVSSGLDREKVEYTLTIQATDMGDGTTTAVAVVEIL 319
Db 309 QDPFLPHKMFVNRTGTIVSVLTSGDRESYTYTLVVOAADLQEGGLSTTAKAVITVK 368
QY 320 DANDNAPMFPDQKYEAHVPENAVGVHEVQRLTVTDLDPAPNSPAWRATYILIMGGDGDGHFTI 379
Db 369 DINDNAPVFNPSYQGVPENEVNARIATLKVTDGDPAPNTPAWKVYTVV-NDPQOQFVV 427
QY 380 THPESNOGILTRKGLDPEAKNQHTLYVEVTNEAPFVLKLTSTATIVVHVVEDVNEAPV 439
Db 428 VTDPTTNDGILTKAGLDPEAKQVYTLHVRVENEPEFEGSLVPSTATVTVVVDVNEAPI 487
QY 440 FVPPSKVVEVQEGIPTEGPEVCVYTAEDPDK-ENOKISYRILRDPAGWLAMDPSDGOVTAV 498
Db 488 FMPAERRVEVPDFGVGQEITSYTAREPDTFMDQKITRYIWRDTANWLSEINPETGAIFTR 547
QY 499 GTLDREDEOFVRNNIYEVNVLAMDNGSPPTTGTGTLTLLTLDVNDHGPVPEPRQITICNQ 558
Db 548 AEMDREDAEHVKNSTYVALIATDDGSPATGTGTLTLLVLDVNDNAPPEPRNQFCQR 607
QY 559 SPVRHVLNITDKLSFHTSPFOAQLTDDSDIYWTAEVNEEG-DTVVLSLKKFLKQDTYDV 617
Db 608 NPQPHIITILDPDLPNTSPFTAELTHGASVNVNTEYNDAAQESLILQPRKOLEIGEYKI 667
QY 618 HLSLDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGF-----ILPVLGAVLALLF 668
Db 668 HLKLDNQNKQDQVTTLDVHVCDCEGTNNC---MKAGIVAAGLQVPAIILGILGILLALLI 724
QY 669 LLLVLLLVKRRKIKPELILPDDTRDNVYVYEGEGGEEDQDYITOLHRLGLEARPEV 728
Db 725 LILLLLLFLRRRTVVEKPLPPDDTRDNVYVYDEEGGEEDQDFLSQLHRLGLDARPEV 784
QY 729 VLRNDVAPTIPTPMYRPRPANDEIGNFIENLKAANTDPTAPPYDTLLVDFYEGSGSD 788
Db 785 T-RNDVAPTLMSVQYRPRPANDEIGNFIENLKAADSPTAPPYDLSLLVDFYEGSGSE 843
QY 789 AASLSSLTSSASDQDQDYDLNFWGSRFKKLADMYGGEDD 829
Db 844 AASLSSLSNESDQDQDYDLNFWGNRFRKKLADMYGGEDD 884

RESULT 6
US-08-472-481-7
; Sequence 7, Application US/08472481
; Patent No. 5863804
; GENERAL INFORMATION:
; APPLICANT: Ranscht, Barbara
; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,481
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/213,361
; FILING DATE: 14-MAY-1994
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 07/607,293
; FILING DATE: 30-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1686
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 884 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-472-481-7

Query Match 54.6%; Score 2384; DB 1; Length 884;
Best Local Similarity 53.5%; Pred. No. 4,1e-213;
Matches 471; Conservative 123; Mismatches 221; Indels 66; Gaps 13;

QY 9 ASLLLLQV-CWLOCAASEP--CRAVFREAEVTLAAGGAEQEPQALGVK-FMGCPEQPEA 64
Db 10 ALLLLQLVSSML-CQELEPESCSPGFSSEVYTFPVERHLERGHVLRVRFEGCTGRPT 68
QY 65 LFSTNDNDFTVRNGTVOERRSLK----- 88
Db 69 AFFESDSRFKATDGTITVYKRLHLKLETSFLVRAROSSHRELSTKVTLKSMGHHHRH 128
QY 89 -----RNP-LKTFPSKRI-LRRHKRDWVAPISVPENKGPPORLNQKSKNDROT 139
Db 129 HHRDPASESNPELLMPPSVYGLRRQKRDWVIPPISCPENKEGFPKVLQVLSNRDKET 188
QY 140 KIFYSITPGADSPPEGVFAVEKETGWLKLLNKPLDREEIAKYELFCHAVSENGASVEDPM 199
Db 189 KVFYSITGQADKPPGVFIIERETGWLKVTPQLDREALAKYILKSHAVSSNGEAVEDEM 248
QY 200 NISIIVTQNDHKKPKTQDTFRGSLVLEGLVPGTSVMQVTTATDEDDDAIYTYNGVAVYSIHS 259
Db 249 EIVITVTDQNDNRPEFTQEVFEGSVAGAVPGTSVMKVSATADDDVNTYNAAIYTVS 308
QY 260 QEPKDPHDLMTFTHRTGTISVSSGLDREKVEYTLTIQATDMGDGTTTAVAVVEIL 319
Db 309 QDPFLPHKMFVNRTGTIVSVLTSGDRESYTYTLVVOAADLQEGGLSTTAKAVITVK 368
QY 320 DANDNAPMFPDQKYEAHVPENAVGVHEVQRLTVTDLDPAPNSPAWRATYILIMGGDGDGHFTI 379
Db 369 DINDNAPVFNPSYQGVPENEVNARIATLKVTDGDPAPNTPAWKVYTVV-NDPQOQFVV 427
QY 380 THPESNOGILTRKGLDPEAKNQHTLYVEVTNEAPFVLKLTSTATIVVHVVEDVNEAPV 439
Db 428 VTDPTTNDGILTKAGLDPEAKQVYTLHVRVENEPEFEGSLVPSTATVTVVVDVNEAPI 487
QY 440 FVPPSKVVEVQEGIPTEGPEVCVYTAEDPDK-ENOKISYRILRDPAGWLAMDPSDGOVTAV 498
Db 488 FMPAERRVEVPDFGVGQEITSYTAREPDTFMDQKITRYIWRDTANWLSEINPETGAIFTR 547
QY 499 GTLDREDEOFVRNNIYEVNVLAMDNGSPPTTGTGTLTLLTLDVNDHGPVPEPRQITICNQ 558
Db 548 AEMDREDAEHVKNSTYVALIATDDGSPATGTGTLTLLVLDVNDNAPPEPRNQFCQR 607
QY 559 SPVRHVLNITDKLSFHTSPFOAQLTDDSDIYWTAEVNEEG-DTVVLSLKKFLKQDTYDV 617
Db 608 NPQPHIITILDPDLPNTSPFTAELTHGASVNVNTEYNDAAQESLILQPRKOLEIGEYKI 667
QY 618 HLSLDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGF-----ILPVLGAVLALLF 668
Db 668 HLKLDNQNKQDQVTTLDVHVCDCEGTNNC---MKAGIVAAGLQVPAIILGILGILLALLI 724
QY 669 LLLVLLLVKRRKIKPELILPDDTRDNVYVYEGEGGEEDQDYITOLHRLGLEARPEV 728
Db 725 LILLLLLFLRRRTVVEKPLPPDDTRDNVYVYDEEGGEEDQDFLSQLHRLGLDARPEV 784
QY 729 VLRNDVAPTIPTPMYRPRPANDEIGNFIENLKAANTDPTAPPYDTLLVDFYEGSGSD 788
```


Db 785 T-RNDVAPLMSVQYRPRANPEIDGNFIDENLKAAADSDPTAPPYDLSLVDFYEGSGSE 843
Qy 789 AASLSSTSSASDQDDYDYLNEWGSRFKKLADMYGGEDD 829
Db 844 AASLSLSNESDQDDYDYLNEWGNRFKKLADMYGGEDD 884

RESULT 7
US-08-237-919-2
; Sequence 2, Application US/08237919
; Patent No. 5610281
; GENERAL INFORMATION:
; APPLICANT: Brenner, Michael B
; APPLICANT: Cepek, Karyn L
; TITLE OF INVENTION: Methods and Compositions for
; Modulating Heterotypic E-cadherin Interactions with T lymphocytes
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,919
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 878 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-237-919-2

Query Match 53.8%; Score 2348.5; DB 1; Length 878;
Best Local Similarity 54.0%; Pred. No. 8.4e-210;
Matches 472; Conservative 115; Mismatches 230; Indels 57; Gaps 12;

Qy 8 LASLLQLQVCMQAASPCRAVREAEVILEAGAEQEPQALGKV-FMCGPQGE--- 62
Db 8 LSGLLLRSLPGSQERSPPCLTRELHVHGAPEKRP--RLGRVNFEDCTGRQRTAI 65
Qy 63 -----PALFSTNDNDFTVRNGETVQERSLK-----ERNP--- 92
Db 66 FLTPPIPKVGTGDIIVTKRPLRPHNPTDPFLGLRWDSTYRKFSKVTLLNTVGHHRPPPHQ 125
Qy 93 -----LKIFP-SKEILRRHKEDWVAPISVPENCKGPPQRLNOLKSKDRTKIF 142
Db 126 ASVSGIQAEILTFPNSPGLRQRKQDWVIPPISCPENENKGPFPKNLVQIRKSKQKEGVF 185
Qy 143 YSITGPAGDSPGEGVAFVEKETGMLLLNKLDRBIEAKYELFHHAVSNGASVEDPMNIS 202
Db 186 YSITQGADTPVGVFIETREITGWLKVTEPLDRERIATYTLFSAVSSNGNAVEDPMEL 245
Qy 203 IIVTDQNDHKPKFTQDTPRGSVLEGLVPGTSVMQVATDEDAIYTYNGVVAISHQEP 262
Db 246 ITVTDQNDNKBEFTQEVFKGSMEGALPGTSVMEVTATDADDVNTVNAAIAYTILSQDP 305

Qy 263 KDPHDLMTFTHRTGTISVISGLDREKVPYETLTITQATMDGDGSGSTTTTAVAVVEILDAN 322
Db 306 ELFPDKNMFNIRNTGIVSVTTTGLDRESPTTLLVVOAADLQGEGLSTTATAVITVTDN 365
Qy 323 DNAPMTDPKQYEAHVPENAVGVHQBRLTVTDLDAPNSPAWRATYLLMGDDGDDHFTTTH 382
Db 366 DNPPIFNPTTYKQVPEANVVTITLKVTDADAPNTPAWEAVYTL-NDGGQGVVVTIN 424
Qy 383 PESNOGILTRKGLDFEAKNQHTLYXEVTNEAPFVLKLPSTSTATIVVHVVDVNEAPVFP 442
Db 425 PVNNDGILKTAKGLDFEAKQYILHVAVTNVPEVSLTSTATVTVDLVDVNEGPVFP 484
Qy 443 PSKVEVQEGIPTGEPVCVYTAEDPPK-ENQKISYRILRDPAGWLAMPDPSQGVAVGTL 501
Db 485 PEKRVESVDFGVGQEITSYTAQEPDTFMEQKITRYRWRDTRNWLINPDTGAISTRAL 544
Qy 502 DRDEQFVRNIIYEVWVLMNDGSPPTGTGTLILLILLIDVNDHGPVPEPQITICNQSPV 561
Db 545 DREDFEHVKNSTYALIIATDNGSPVATGTGTLILLISLVNDNAPPEPRTIFFCERNPK 604
Qy 562 RVLNITDKDLSPTSPFOALTDDSDIYVTAEVNE-EGDTVVLKFLKQDTYDVHLS 620
Db 605 PQVINIHADLPNTSPFTAEALTHGRVPNWTIQVNDPTQESIILKPKMALEVGDYKINLK 664
Qy 621 LSDHGNKEQLTVIRATVCDCHGHVETC--PGPWKGSF----ILPVLGAVLALLFLLLVLL 674
Db 665 LMDNQKQVTTLEVSVCDCGGAAGVCRKAQPVAGLQIPAILGILGILALLLILL 724
Qy 675 LTVRKRRKIKEPILLPEDDTRDNVFFYGBGGGEEQDDYDITQLHRLGLEARPEVVLNDV 734
Db 725 LFLRRRAVVKPELLPPEDDTRDNVFFYGBGGGEEQDDYDITQLHRLGLEARPEV- 783
Qy 735 APTIIPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDITLLVDFYEGSGDAASLS 794
Db 784 APTLMSVRYLPRPANPDEIGNFIDENLKAANTDPTAPPYDITLLVDFYEGSGDAASLS 843
Qy 795 LTSASDQDDYDYLNEWGSRFKKLADMYGGED 828
Db 844 LNSSESXKDDYDYLNEWGNRFKKLADMYGGED 877

RESULT 8
US-08-732-429-2
; Sequence 2, Application US/08732429
; Patent No. 6300080
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Methods and Compositions for Modulating Heterotypic
; E-cadherin Interactions with T Lymphocytes
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/732,429
; FILING DATE: herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/237,919
; FILING DATE: 3 May 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7023
; TELECOMMUNICATION INFORMATION:


```
Db 605 PQVINIHADLPNTSPFTAEHLTHGRVPMNTIQNDPTQESIIILKPKMALEVGDKINLK 664
Qy 621 LSDHGNKEQLTVIRATVCDCHGHVETC--PGPMKGGF-----ILPVLGAVLALLFLILVLL 674
Db 665 LMDNQNDQVTTLEVSVCDEGAAGVCRKAQPVAGLQIPAILGILGGIALLILILLLL 724
Qy 675 LLVRKKRRIKEPLLLPEDDTRDNVYFYEGEGGEDQDYDITQIHLRGLEARPEVVLKNDV 734
Db 725 LFLRRRAVKEPLLPEDDTRDNVYFYDEEGGEDQDFDLSQLHGLDARPEVT-RNDV 783
Qy 735 APTIIPMYRPRANPDEIGNFIENLKAANTDPTAPPYDITLLVDFYEGSGDAASLSS 794
Db 784 APTLMSVRYLPRANPDEIGNFIENLKAADTPTAPPYDITLLVDFYEGSGEASLSS 843
Qy 795 LTSASDQDDYDYLNEGWSRPFKCLADMYGGGD 828
Db 844 LNSSESDDQDDYDYLNEGWSRPFKCLADMYGGGD 877

RESULT 10
US-09-798-267-3
; Sequence 3, Application US/09798267
; Patent No. 6406870
; GENERAL INFORMATION:
; APPLICANT: Brenner, Michael
; APPLICANT: Cepek, Karyn
; TITLE OF INVENTION: Methods and Compositions for Modulating Heterotypic E-cadherin
; FILE OF INVENTION: Interactions with T Lymphocytes
; FILE REFERENCE: L0560/7008RP
; CURRENT APPLICATION NUMBER: US/09/798,267
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 08/237,919
; PRIOR FILING DATE: 1994-05-03
; PRIOR APPLICATION NUMBER: PCT/US 95/05518
; PRIOR FILING DATE: 1995-05-03
; PRIOR APPLICATION NUMBER: US 08/732,429
; PRIOR FILING DATE: 1996-11-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 878
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-798-267-3

Query Match 53.8%; Score 2348.5; DB 2; Length 878;
Best Local Similarity 54.0%; Pred. No. 8.4e-210;
Matches 472; Conservative 115; Mismatches 230; Indels 57; Gaps 12;

Qy 8 LASLILLOVCHLQCAASEPCRAVREAEVTLLEAGGAEQEPQALGKV-FMCCPQGE---- 62
Db 8 LSGLLLLRLSLGQERSPPCLTRHLVHGAPPEKRP--RLGRVNFEDCTGRQRTAI 65
Qy 63 -----PALFSTDNDDFTVRNGEFTVQERRSLK-----ERNP--- 92
Db 66 FLTPIPKVGTGVTIVKRLPHNPTDPFLGLRWDSTYRKSTKVTNTVGHHRPPPHQ 125
Qy 93 -----LKIFF-SKRLRRHRKDWVAVIPSPENGKGPFPQRLNQLKSNKDRDRTKIF 142
Db 126 ASVSGIQAEALLTFPNSPGLRRQKRDVWIPISCPENEGKGFPPKLVQIKSNKDKGKVF 185
Qy 143 YSITGPGADSPBEGVFAVEKETGMLLNKPLDRREIAKYELFHAVSENGASVEDPMNIS 202
Db 186 YSITGGQADTFPVGVFIERTETGLVKVTEPLDRERIATYTLFHAVSSNGNAVEDPMEL 245
Qy 203 IIVTDQNDHKPKFQDTPRGSVLEGLVPGTSMQVMTATDEDAIYTVNGVAVYIHSQEP 262
Db 246 ITVTDQNDKPEFTQEVFKGSMEGALPCTSVMEVMTATDADDVNTYNAIAYTILSQDP 305
Qy 263 KPDHLMPTIHRSTGTISVISGLDREKVPYTTITQATMDMDGSGSTTAVAVVEILDAN 322
Db 306 ELDPKNMFTINRNTGVISVWTTGLDRSFYTYTLVQQAADLQGEGLSTTATAVITVDTN 365
```

```
Qy 323 DNAPMFDPQYEAHVPENAVGHEVQRLTVTDLDPNPSPAWRATYLLIMGDDGDHFTITTH 382
Db 366 DNPPIFNPTTYKQOVENEANVITLTKVTDADAPNTPAWEAVYTIL-NDGQGFVVTTN 424
Qy 383 PESNOGILTRKGLDPEAKNQHTLYVEVNEAPFVILKLPSTSTATIVVHVEDVNEAPVFP 442
Db 425 PVNNDGILKTAKGLDPEAKQOYLTHVAVTNVVPEVSLTSTATVTVLDVNEGPVFP 484
Qy 443 PSKVEVQEGIPGEPVCVVTAEEDPK-ENOKLSYRILRDPAGWLAMDPSGOVAVGTL 501
Db 485 PEKREVESEDFGVQSEITSYTAQEPDTFMEQKITRWRNRWLEINPTGAISTRABL 544
Qy 502 DRDEQFVRNIVYEVWVLAMDNGSPPTGTGTLTLLTLDVNDHGPVPEPRQITICNOSP 561
Db 545 DREDFEHVKNSTYALIATDNGSPVATGTGTLTLLTLDVNDNAPPEPRPTIFFCERNPK 604
Qy 562 RHLNITDKDLSPTSPFQAQLTDDSDIYWAEBNE-EGTVTVLSLKKFKLQDITYVHLS 620
Db 605 PQVINIHADLPNTSPFTAEHLTHGRVPMNTIQNDPTQESIIILKPKMALEVGDKINLK 664
Qy 621 LSDHGNKEQLTVIRATVCDCHGHVETC--PGPMKGGF-----ILPVLGAVLALLFLILVLL 674
Db 665 LMDNQNDQVTTLEVSVCDEGAAGVCRKAQPVAGLQIPAILGILGGIALLILILLLL 724
Qy 675 LLVRKKRRIKEPLLLPEDDTRDNVYFYEGEGGEDQDYDITQIHLRGLEARPEVVLKNDV 734
Db 725 LFLRRRAVKEPLLPEDDTRDNVYFYDEEGGEDQDFDLSQLHGLDARPEVT-RNDV 783
Qy 735 APTIIPMYRPRANPDEIGNFIENLKAANTDPTAPPYDITLLVDFYEGSGDAASLSS 794
Db 784 APTLMSVRYLPRANPDEIGNFIENLKAADTPTAPPYDITLLVDFYEGSGEASLSS 843
Qy 795 LTSASDQDDYDYLNEGWSRPFKCLADMYGGGD 828
Db 844 LNSSESDDQDDYDYLNEGWSRPFKCLADMYGGGD 877

RESULT 11
PCT-US95-05518-2
; Sequence 2, Application PC/TUS9505518
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Methods and Compositions for Modulating Heterotypic
; TITLE OF INVENTION: E-cadherin Interactions with T Lymphocytes
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05518
; FILING DATE: herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/237,919
; FILING DATE: 3 May 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
```

		; Sequence 7, Application US/08474067	
		; Patent No. 5811518	
		; GENERAL INFORMATION:	
		; APPLICANT: Ranscht, Barbara	
		; TITLE OF INVENTION: T-Cadherin Adhesion Molecule	
		; NUMBER OF SEQUENCES: 9	
		; CORRESPONDENCE ADDRESS:	
		; ADDRESSEE: Campbell and Flores	
		; STREET: 4370 La Jolla Village Drive, Suite 700	
		; CITY: San Diego	
		; STATE: California	
		; COUNTRY: United States	
		; ZIP: 92122	
		; COMPUTER READABLE FORM:	
		; MEDIUM TYPE: Floppy disk	
		; COMPUTER: IBM PC compatible	
		; OPERATING SYSTEM: PC-DOS/MS-DOS	
		; SOFTWARE: Patent In Release #1.0, Version #1.25	
		; CURRENT APPLICATION DATA:	
		; APPLICATION NUMBER: US/08/474,067	
		; FILING DATE: 07-JUN-1995	
		; CLASSIFICATION: 514	
		; PRIOR APPLICATION DATA:	
		; APPLICATION NUMBER: US 08/213,361	
		; FILING DATE: 14-MAY-1994	
		; PRIOR APPLICATION DATA:	
		; APPLICATION NUMBER: US 07/607,293	
		; FILING DATE: 30-OCT-1990	
		; ATTORNEY/AGENT INFORMATION:	
		; NAME: Campbell, Cathryn A.	
		; REGISTRATION NUMBER: 31,815	
		; REFERENCE/DOCKET NUMBER: P-LJ 1682	
		; TELECOMMUNICATION INFORMATION:	
		; TELEPHONE: (619) 535-9001	
		; TELEFAX: (619) 535-8949	
		; INFORMATION FOR SEQ ID NO: 7:	
		; SEQUENCE CHARACTERISTICS:	
		; LENGTH: 837 amino acids	
		; TYPE: amino acid	
		; TOPOLOGY: linear	
		US-08-474-067-7	
		Query Match 51.1%; Score 2231.5; DB 1; Length 837;	
		Best Local Similarity 54.0%; Pred. No. 6.6e-199;	
		Matches 450; Conservative 117; Mismatches 212; Indels 55; Gaps 13;	
QY	48	GOALGKV-FMCGCPGEPALF-STD-----NDD-----FTVRNGSTVOE 83	
DB	6	GRELGRVFAACGRPWAVYPTDTRFKVNGGVVSTKRPLTYGRKISFTIYAQDAMGK 65	
QY	84	RRS-----LKERNP LKI-PP--SKRILRRHRKRDVWVAPISVPENKKG 122	
DB	66	RHSARVTVGRHRRHRRHHNHQLDITTPAVLTTPKHDPGFLRRQKRDWVIPPISCLNHRG 125	
QY	123	PPQRLNQLKSNKDRDKIFYSITGPGADSPPEGVAVVEKETGWLKLNKPLDREETAKYE 182	
DB	126	PYEMRLVQIKSNKDKESKVVYSITGQADSPVPGFIIERETGWLSEVTLQDLREKIDRYT 185	
QY	183	LFEGHVAENGASVDEPMNISILVTQNDHKPKFTODTRGSLGVLPGTSMOVTATDE 242	
DB	186	LUSHVASAGQVQVDEPMELIITVMDQNDKNPFIIEFVFGVYIEENAKPGTSMVTVNATDA 245	
QY	243	DDAIYTYGVAVYSHSQEPKDPHLMFTIHRSTGTISVISGLDREKPYEYTLTIQATD 302	
DB	246	DDAVNTDNGIVSYISVSQPPRPHQMFITIDPAKGIISVLGTGLDRETTNPVTLIVQATD 305	
QY	303	MDGDSSTTAVAVVEILDANDNAPDPDPKYIAHPENAVGHEVQRLTYTDLDPAPNSPAW 362	
DB	306	QEGKGLSNATATAIIEVTDANDNIPFNPTMYEGVVEENKPGTEVARLTVDQDAPGSPAW 365	
QY	363	RATYILMGDDGDGHFTITTHPESNQGILTRKGLDPEAKNQHTLVYVEVTNEAPFVLKLP 422	
DB	366	QAVYHIKSGNLGAFSIITDPSTNNGILTKAGLDYETKSRYDLVVTVENKPLSPVITL 425	


```
Db      364 NTATAVITVTDVNDNPPEFTAMTFYGEVPENRVDIIIVANLTVTDKQOPHTPAWNAVYRIS 423
Qy      370 GGDGDHFTITTHPESNOGLTTRKGLDPEAKQOHTLYVEVTNEAPFVLXL---PTSTAT 426
Db      424 GGDPTGRFAJQTDNSNDGLVTVVKPIDFETNRMFVLTVAENQVPLAKGIOHPQSTAT 483
Qy      427 IVHVEDVNEAPVFPVPSKVVEQVGEPIGTGEPVCVYTAEDPK-ENQKISYRIIRDPAW 485
Db      484 VSVTVIDVNEPYPAPNPKIIRQEEGLHAGTMLTFTTAQDPDRYMQQNIYTKLSDFANW 543
Qy      486 LAMPDGSQVAVGTLDRDEQFVRNNIYVMVLAMDNQSPPTGTCTLLTLIDVNDHG 545
Db      544 LKIDPVNGQITTAVLAVLDRSPN-VKNNIYNATFLASDNQIPPMSTGTCTLQIYLLIDINDNA 602
Qy      546 PVPEPRQITICNQSPVRHVLNIT--DKDLSPHTSFFQAQLTDDSDIY---WTAEVNEEGD 600
Db      603 FQVLPQAEATC-ETPDPNSINITALDYDIDFNAGFFAFDLPLSPVTIKRNTI--TRLNGD 660
Qy      601 TVVLSLK-KFLKQDQYDVHLSLSLSDHGN--KEQLTVIRATVCDCHG-----VETCPGFWK 652
Db      661 FAQLNLKIKFEAGIYEVPIIITDSGNPPKSNISILAVKYQCDSNGDCTDQDRIVGAGL 720
Qy      653 GGFILPVLGAVLALLFLLVLLLV-----RKRRKIKEPLLLPEDDTRDNVYYG 702
Db      721 G-----TGAIIATLLCIIILLILVLMFVVMKRRDKERQAKQLLIDPEDDVRDNILKYD 774
Qy      703 BEGGGEEDQDYDITQLHRGLEARPEV-----VLRNDVAPTIIPTPMYRPRPA--NPDEIG 755
Db      775 BEGGGEEDQDYDLSLOQOPDTVEPDAIKPVGIRMDERP-IHAEPQYFVRSAAHPGDI 833
Qy      756 NFIIENLKAANTDPTAPPYDTLVFDYEGSGSDAASLSSLTSSASDQDQDYDYLNEWGSR 815
Db      834 DFINEGLKAANDPTAPPYDLSLVFDYEGSGSTAGSLSSLNSSSSGGEQDYDYLNDWGP 893
Qy      816 FKLADMYGGGED 828
Db      894 FKLADMYGGGDD 906
```

Search completed: December 12, 2005, 09:17:05
Job time : 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2005, 09:11:23 ; Search time 11 Seconds
(without alignments)
420.863 Million cell updates/sec

Title: US-10-696-639-39

Perfect score: 4369

Sequence: 1 MGLPRGPLASLLLLQVCMWLQ.....NEWGRFKLADMYGGEDD 829

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 32527 seqs, 5584426 residues

Total number of hits satisfying chosen parameters: 32527

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.New.*

- 1: /cgn2_6/prodata/1/pubpa/US09_NEW_PUB.pap.*
- 2: /cgn2_6/prodata/1/pubpa/US06_NEW_PUB.pap.*
- 3: /cgn2_6/prodata/1/pubpa/US07_NEW_PUB.pap.*
- 4: /cgn2_6/prodata/1/pubpa/US08_NEW_PUB.pap.*
- 5: /cgn2_6/prodata/1/pubpa/PCT_NEW_PUB.pap.*
- 6: /cgn2_6/prodata/1/pubpa/US10_NEW_PUB.pap.*
- 7: /cgn2_6/prodata/1/pubpa/US11_NEW_PUB.pap.*
- 8: /cgn2_6/prodata/1/pubpa/US60_NEW_PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4361	99.8	829	7	US-11-090-739-122 Sequence 122, App
2	4361	99.8	829	7	US-11-186-284-18 Sequence 18, Appl
3	982	22.5	790	7	US-11-080-991-82 Sequence 82, Appl
4	957	21.9	896	7	US-11-080-991-98 Sequence 98, Appl
5	919	21.0	801	7	US-11-174-150-29 Sequence 29, Appl
6	796	18.2	824	6	US-10-821-234-1008 Sequence 1008, Ap
7	588	13.5	589	7	US-11-174-150-28 Sequence 28, Appl
8	526.5	12.1	832	7	US-11-108-172-1081 Sequence 1081, Ap
9	526.5	12.1	958	7	US-11-108-172-1087 Sequence 1087, Ap
10	408.5	9.3	1060	7	US-11-090-739-120 Sequence 120, App
11	381.5	8.7	807	7	US-11-102-240-98 Sequence 98, Appl
12	333.5	7.6	1184	6	US-10-131-826A-394 Sequence 394, App
13	280	6.4	530	6	US-10-131-826A-130 Sequence 130, App
14	240.5	5.5	1730	7	US-11-192-967-4 Sequence 4, Appli
15	240.5	5.5	1730	7	US-11-193-715-4 Sequence 4, Appli
16	226	5.2	72	6	US-10-509-170-19 Sequence 19, Appl
17	214	4.9	1734	7	US-11-192-967-6 Sequence 6, Appli
18	214	4.9	1734	7	US-11-193-715-6 Sequence 6, Appli
19	207	4.7	1717	7	US-11-192-967-2 Sequence 2, Appli
20	207	4.7	1717	7	US-11-193-715-2 Sequence 2, Appli
21	147.5	3.4	2384	6	US-10-821-234-1545 Sequence 1545, Ap
22	126.5	2.9	1346	7	US-11-060-005-2 Sequence 2, Appli
23	125.5	2.9	2647	6	US-10-821-234-1303 Sequence 1303, Ap
24	124.5	2.8	895	7	US-11-150-406-2 Sequence 2, Appli
25	124.5	2.8	920	6	US-10-821-234-1129 Sequence 1129, Ap

RESULT 1

US-11-090-739-122
; Sequence 122, Application US/11090739
; Publication No. US20050260639A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Yusuke
; APPLICANT: KATAGIRI, Toyomasa
; APPLICANT: NAKAGAWA, Hidewaki
; TITLE OF INVENTION: METHOD FOR DIAGNOSING PANCREATIC CANCER
; FILE REFERENCE: 082368-003600US
; CURRENT APPLICATION NUMBER: US/11/090,739
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: PCT/JP2003/011817
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: US 60/555,809
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: US 60/450,889
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/414,872
; PRIOR FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 122
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-739-122

Query Match 99.8%; Score 4361; DB 7; Length 829;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 828; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MGLPRGPLASLLLLQVCMWLQCAASPCRAVFREAEVTLFAGGAEQEPGQALGVFMGCPG	60
Db	1	MGLPRGPLASLLLLQVCMWLQCAASPCRAVFREAEVTLFAGGAEQEPGQALGVFMGCPG	60
Qy	61	QSPALFSTDNDFTVRNGETVQERRSLKERNPLKIPPSKRIILRRHKRDVWVAPISVPENG	120
Db	61	QSPALFSTDNDFTVRNGETVQERRSLKERNPLKIPPSKRIILRRHKRDVWVAPISVPENG	120
Qy	121	KGPPFQRLNQLKSNKDRDTKIFYSITGPGADSPPGVFAVEKETGWLILNKPLDREEIAK	180
Db	121	KGPPFQRLNQLKSNKDRDTKIFYSITGPGADSPPGVFAVEKETGWLILNKPLDREEIAK	180
Qy	181	YELFGHVSNGASVEDPMNISIIITVDQNDHKPKFTQDTFRGSLVLEGLVPGTSMQVTAT	240
Db	181	YELFGHVSNGASVEDPMNISIIITVDQNDHKPKFTQDTFRGSLVLEGLVPGTSMQVTAT	240

ALIGNMENTS

26	123.5	2.8	826	6	US-10-793-626-1066	Sequence 1066, Ap
27	121	2.8	428	6	US-10-689-742-50	Sequence 50, Appl
28	118.5	2.7	559	6	US-10-793-626-1376	Sequence 1376, Ap
29	118	2.7	1596	7	US-11-060-005-4	Sequence 4, Appli
30	107.5	2.5	616	7	US-11-090-878-20	Sequence 20, Appl
31	107.5	2.5	624	7	US-11-090-878-2	Sequence 2, Appli
32	107.5	2.5	637	7	US-11-090-878-24	Sequence 24, Appl
33	106.5	2.4	596	7	US-11-090-878-22	Sequence 22, Appl
34	106.5	2.4	617	7	US-11-090-878-18	Sequence 18, Appl
35	106.5	2.4	3063	7	US-11-186-284-26	Sequence 26, Appl
36	105.5	2.4	5024	6	US-10-793-626-2964	Sequence 2964, Ap
37	105	2.4	720	6	US-10-793-626-2058	Sequence 2058, Ap
38	105	2.4	1897	6	US-10-821-234-1635	Sequence 1635, Ap
39	105	2.4	1907	7	US-11-000-463-250	Sequence 250, App
40	104	2.4	7968	7	US-11-186-731-5	Sequence 5, Appli
41	103.5	2.4	598	7	US-11-090-878-8	Sequence 8, Appli
42	103.5	2.4	910	6	US-10-131-826A-112	Sequence 112, App
43	103.5	2.4	1141	6	US-10-601-368-24	Sequence 24, Appl
44	103.5	2.4	1166	6	US-10-601-368-22	Sequence 22, Appl
45	103.5	2.4	1188	6	US-10-601-368-21	Sequence 21, Appl


```
Qy 241 DEDDAIYTYGVVAYSIIHQEPKDPHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
Db 241 DEDDAIYTYGVVAYSIIHQEPKDPHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
Qy 301 TMDGSGSTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVEQRLTVDLDAPNSP 360
Db 301 TMDGSGSTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVEQRLTVDLDAPNSP 360
Qy 361 AWRTYILMGDDGDHFTIITHPESNQILTTKGLDFEAKNQHTLVVEVNEAPFVLKL 420
Db 361 AWRTYILMGDDGDHFTIITHPESNQILTTKGLDFEAKNQHTLVVEVNEAPFVLKL 420
Qy 421 PTSTATIIVHVEDVNEAPFVPPSKVVEQEGIPTGEPVCVYTAEDDPKENQKISYRILR 480
Db 421 PTSTATIIVHVEDVNEAPFVPPSKVVEQEGIPTGEPVCVYTAEDDPKENQKISYRILR 480
Qy 481 DPAGWLAMPDSGQVTAAGTLDREDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLLID 540
Db 481 DPAGWLAMPDSGQVTAAGTLDREDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLLID 540
Qy 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSLKKFLKQDYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Db 601 TVVLSLKKFLKQDYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Qy 661 GAVLALLFLLVLLVLRKKRKIKEPILLPDDTRDNVYFYGGEGGEEQDYDITQLHR 720
Db 661 GAVLALLFLLVLLVLRKKRKIKEPILLPDDTRDNVYFYGGEGGEEQDYDITQLHR 720
Qy 721 GLEARPEVLRNDVAPTIIPPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVLRNDVAPTIIPPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
Qy 781 DYEGSGSDAASLSLTSASDQDQDYVLYNEWGSFRFKKLADMYGGGDD 829
Db 781 DYEGSGSDAASLSLTSASDQDQDYVLYNEWGSFRFKKLADMYGGGDD 829

RESULT 2
US-11-186-284-18
; Sequence 18, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kanatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MP001-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; PRIOR FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 829
; TYPE: PRT
```

```
; ORGANISM: Homo Sapiens
US-11-186-284-18
Query Match 99.8%; Score 4361; DB 7; Length 829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 828; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MGLPRGPLASLLQLQVLCWQCAASEPCRAVFRAEVTLEAGGAEQPGQALGVFVGCPCG 60
Db 1 MGLPRGPLASLLQLQVLCWQCAASEPCRAVFRAEVTLEAGGAEQPGQALGVFVGCPCG 60
Qy 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDVWVAIPSPENG 120
Db 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDVWVAIPSPENG 120
Qy 121 KGPPORLNLKSNKORDTKIFYSIITGPGADSPPEGVFAVEKETGWLKLLNKLDPREIAK 180
Db 121 KGPPORLNLKSNKORDTKIFYSIITGPGADSPPEGVFAVEKETGWLKLLNKLDPREIAK 180
Qy 181 YELFGHAVSENGASVEDPMNISIITVDQNDHKPKFTQDTPRGVLEGVLPGTSMQVAT 240
Db 181 YELFGHAVSENGASVEDPMNISIITVDQNDHKPKFTQDTPRGVLEGVLPGTSMQVAT 240
Qy 241 DEDDAIYTYGVVAYSIIHQEPKDPHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
Db 241 DEDDAIYTYGVVAYSIIHQEPKDPHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
Qy 301 TMDGSGSTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVEQRLTVDLDAPNSP 360
Db 301 TMDGSGSTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVEQRLTVDLDAPNSP 360
Qy 361 AWRTYILMGDDGDHFTIITHPESNQILTTKGLDFEAKNQHTLVVEVNEAPFVLKL 420
Db 361 AWRTYILMGDDGDHFTIITHPESNQILTTKGLDFEAKNQHTLVVEVNEAPFVLKL 420
Qy 421 PTSTATIIVHVEDVNEAPFVPPSKVVEQEGIPTGEPVCVYTAEDDPKENQKISYRILR 480
Db 421 PTSTATIIVHVEDVNEAPFVPPSKVVEQEGIPTGEPVCVYTAEDDPKENQKISYRILR 480
Qy 481 DPAGWLAMPDSGQVTAAGTLDREDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLLID 540
Db 481 DPAGWLAMPDSGQVTAAGTLDREDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLLID 540
Qy 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSLKKFLKQDYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Db 601 TVVLSLKKFLKQDYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Qy 661 GAVLALLFLLVLLVLRKKRKIKEPILLPDDTRDNVYFYGGEGGEEQDYDITQLHR 720
Db 661 GAVLALLFLLVLLVLRKKRKIKEPILLPDDTRDNVYFYGGEGGEEQDYDITQLHR 720
Qy 721 GLEARPEVLRNDVAPTIIPPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVLRNDVAPTIIPPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
Qy 781 DYEGSGSDAASLSLTSASDQDQDYVLYNEWGSFRFKKLADMYGGGDD 829
Db 781 DYEGSGSDAASLSLTSASDQDQDYVLYNEWGSFRFKKLADMYGGGDD 829

RESULT 3
US-11-080-991-82
; Sequence 82, Application US/11080991
; Publication No. US20050266437A1
; GENERAL INFORMATION:
; APPLICANT: Velby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
```



```
Db 714 VFGATGKR-----FPEDIAQONLIISNTEAPGDDRVCSANGFMTQTNNSSQGFCGT 766
Qy 726 PEVLRNDVAPTIIPTMYRPRANPDEI---GNFIENLKAAN-----766
Db 767 MSGMKNGGQETI-----EMKGGNQTLSCRGAGHHHTLDCRGGHTEVD 812
Qy 767 -----TDP-----TAPPYDTLLVFDYEGSGDAASLSLTSAS 800
Db 813 NCRITYSEHSFTOPRLGEKLRHCNQNEDRMPSQDYVLTYNREGSGPAGSVG---CCSEK 870
Qy 801 DQDQDYDYLNEWGRFKKLAD 821
Db 871 QEEDGLDFLNNLEPKFITLAE 891

RESULT 5
US-11-174-150-29
; Sequence 29, Application US/11174150
; Publication No. US20050260714A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall F.
; APPLICANT: Xiang, ZhaoYing
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50022
; CURRENT APPLICATION NUMBER: US/11/174,150
; CURRENT FILING DATE: 2005-07-01
; PRIOR APPLICATION NUMBER: US/10/257,174
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: PCT/US01/11797
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/196,603
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/199,417
; PRIOR FILING DATE: 2000-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-174-150-29

Query Match 21.0%; Score 919; DB 7; Length 801;
Best Local Similarity 33.5%; Pred. No. 3.8e-62;
Matches 255; Conservative 128; Mismatches 322; Indels 56; Gaps 24;

Qy 103 RRHKRDWVAVPISVPENKGPFPQRLNOLKSNKDR-DTKIFYSITGPGADSPPEGVFAVE 161
Db 55 QTRKRWVWNOFFVLEEYTGTDPLVYGLKHSMDRGDSIKYILSGEGAGI-----VFTID 110
Qy 162 KETGMLLNKPLDREEIAKYELFGHAV-ENGASVEDPMNISIIIVTDQNDHKPKFTQDTF 220
Db 111 DTTGDIHAIQLDREERAQYTLRAQALDRRTGRPEPESEFIKIQDINDNEPKFLDGPY 170
Qy 221 RGSVLEGLVPGTSVMQVATDEDAIYTYNGVWAYS IHSQBPKPDLMLFTIHRSTGTIS 280
Db 171 VATVPMSFVGTSVIQVATDADDPYTGNSARVVSILQGP-----YFSVDSKGTGVR 224
Qy 281 VISSGLDREKVPXYTLTIQATDMGD--GSTTTAVAVVEIILDANDNAMPDPQKYEAVHP 338
Db 225 TALMMMDREAKYEYEVIIQAKDMGQLGLAGTTVNTLSDVNDNPPRFPQKHQMSVL 284
Qy 339 ENA-VGHEVQRLTVTDLAPNSPAWRATYILMGDGDGHFTITTHPESNQILTRKGLD 397
Db 285 ESAPISSTVGRVFAKDLDEGINAEK--YTIVDGDAFADISTDPNFQVGIITVKFELS 342
Qy 398 FEAKNQHTLYVEVTN---EAPFVLKLP-TSTATIVVHVVEDVNEAPVFPVPSKVEVQEGI 453
Db 343 FESKKSYYTLKVEGANPHLEMRFLNLGPFQDTTTHVISVEDVDEPVPFPGFYFVEVPEDV 402

454 PTGEPVCVYTAEDPDKENOKISYRIILR--DPAGWLAMDPSGQVTAVTGLDREDEQFVRN 511
Db 403 AIGTTIIISAKDPDVTNNISIRYSIDRSSDPGRFFVVDITTGALMTARPLDREFSW--- 459
Qy 512 NIYEVWVLAMDNQSPPTTGTGTLTLLLDVNDHGP-VPEPRQITICNQSPVRHVL---NI 567
Db 460 --HNITVLAMENMNPQVGSVPVTIKVLVDNDNAPEFFRFEAFVCENAKAGQLIQTVSA 517
Qy 568 TKDLSLPHTSFPQAQITDD--SDIYTAENVESGDTVVLSSKKFLKQDITDV-HLS--LS 622
Db 518 VDODDPRNGOHFYSLAPEAANNPNFTIRDQNONTARILTRRSGRFOEQESVPHLPILIA 577
Qy 623 DHG-----NKEQLTVIRATVCDCHGVETCTCPGWKGFILPV-----LGAVLALLFLLV 672
Db 578 DSGQPVLSSTGTTLTIQVCSDDDDGHVMSCS---PEAYMLPVLSRGLAIIALACIFVLV 634
Qy 673 LLLLV-RKKRKIKPELLL-PEDTRONVIFYGEGGEEEDQD-YDITQLHRGLEAREPV 729
Db 635 LVLLILSMRRHRKQPYIIDDEENIHENIVRYDDEGGEEDEAFDIAAMWNPREAQAGAA 694
Qy 730 --LRNDVAPTIIPTMYRPRP-ANPDEIGNFIENLKAANTDPTAPPYDTLLVFDYEGSG 786
Db 695 PKTRQDMLPEIESLSRYVPQTCANVSTVHSYVLAKLYEADMDLWAPFDSLOTYMFEGDG 754
Qy 787 SDAASLSLTSASDQDDQDYDYLNEWGRFKKLADMYGGE 827
Db 755 SVAGSLSSLOSATSSEQSDFELTDWGPFRKLAELYGASE 795

RESULT 6
US-10-821-234-1008
; Sequence 1008, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeciampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: Pct_SEQ_genes Version 1.0
; SEQ ID NO 1008
; LENGTH: 824
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1008

Query Match 18.2%; Score 796; DB 6; Length 824;
Best Local Similarity 32.1%; Pred. No. 8.2e-53;
Matches 251; Conservative 124; Mismatches 323; Indels 84; Gaps 33;

Qy 89 ERNPLKIFPSKRILRRHKRDWVAVPISVPENKGPFPQRLNOLKSNKDRDKIFVSTGP 148
Db 72 QRDTHSLPLPH---RRKRDMWINQMHIDEKNTSLPHHVGVKIKSSVSR-KNAYLLKGE 127
Qy 149 GADSPPEGVFAVEKETGMLLNKPLDREEIAKYELFGHAV-SENGASVEDPMNISIIIVTD 207
Db 128 YVGK-----VFRVDAETGDVFAIERLDRENISEYHLTAVVDKOTGENLETSPSTIKVHD 183
Qy 208 QNDHKPKFTQDTFRGSVLEGLVPGTSVMQVATDEDAIYTYNGVWAYS IHSQBPKPDL 267
Db 184 VNDNWPVFTHRLENASVPSSAVGTSVISVAVDADDPYTGDAHASVMYQLKGE----- 238
Qy 268 LMFTIHRSTGTISVISSGLDREKVPXYTLTIQATDMGD--DGSTTTAVAVVEIILDANDNAP 326
Db 239 -YFALDMS-GRITITIKSLDREKQARVEIIVVEARDAQGLRDSGTATVLTQINDNFP 296
Qy 327 MFDQPKYEAVHPENA-VGHEVQRLTVTDLAPNSPAWRAT-YLIMGDGDGHFTITTHPE 384
```

Db 297 FFTQTKYTFVVPEDTRVGTSGVSLFVED---PDEPQNRMTKYSLRGDYQDAFTIETNFA 353
QY 385 SNOGILTRKGLDFEAKNOHTLVVEVTN---EAPFVLKLPSTATVWVHVEDNEAPVFW 441
Db 354 HNEGILKPKMLDYEQVIFVEATDPTIDIRYMSPPAGNRAQVIINITDVEEPIFQ 413
QY 442 PPSKVVEQSIGITGEP-VCVYTAEDPKENQKISYRILDRPAGWLAMPDSQQ---VTA 497
Db 414 QPFYHFLKEN--QKKPLIGTVLAMPDAAARHSIGYSIRR-----TSDKGQFFRVTK 463
QY 498 VGTLDREDEQFVRNNIYEVWVLANMGSPPTTGTGT-----LLTLTDVNDHGP-VPE 549
Db 464 KGDYINEKE--LDREVYPWNLTVAEAKELDSTGTPTGKESIVQVHIEVLNDENDNAPEFAK 521
QY 550 PROITTC-NQSPVRHVLNIT--DKDLSPHTSPTFOAOLTDSDIYMTVAEVEEEDTIVVLSL 606
Db 522 PYQPKVCENAVHQLVLOISAIKDKITPRNVKFKFLNTENNFTLT---DNHNTANITV 578
QY 607 K--KFLKQDITYDVH---LSLSHDG--NKEQLTVIRATVCDCH--GHVETCP--GPWKGGF 655
Db 579 KYGQFDREHT-KVHFPLVVISDNGMPSRTGTSTLTVAACKNCQGBFTFCEDMAAQVGS 637
QY 656 ILPVLCVALLFLLVLLVRKKKIKIKEPLI-----LPBDDTRDNVFFYGEGBGEED- 710
Db 638 IQAWAAILLCILITITVITLLIFLRRRLKQARAHKGSVPE--THEQLVTVDESGGEMDT 695
QY 711 QDYDITDQLH---RGLERPEVLRNDVAPTIITPVMYRPPA-----NPDEIGNFIENL 762
Db 696 TSDVSVLNVRRGGAKPPRAL--DARPSLYAQVQKPPRHAPGAGGCGEAMAEVKK 753
QY 763 KAANTDPTAPPYDTLAVFDYEGSGDAASLSLTSASDQDQDYDYLNEWGSFCKLADM 822
Db 754 DEADHDGDPYDTHIYGYEGSESIAESLSLGTSDSDSDVDYDFLNDWGRFKMLAEL 813
QY 823 YG 824
Db 814 YG 815

RESULT 7

US-11-174-150-28
; Sequence 28, Application US/11174150
; Publication No. US20050260714A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall P.
; APPLICANT: Xiang, Zhaoxing
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50022
; CURRENT APPLICATION NUMBER: US/11/174,150
; CURRENT FILING DATE: 2005-07-01
; PRIOR APPLICATION NUMBER: US/10/257,174
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: PCT/US01/11797
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/196,603
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/199,417
; PRIOR FILING DATE: 2000-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-174-150-28

Query Match 13.5%; Score 588; DB 7; Length 589;
Best Local Similarity 34.1%; Pred. No. 2.9e-37;
Matches 165; Conservative 72; Mismatches 215; Indels 32; Gaps 13;

QY 103 RRHKRDWVAPISVPENGKGPFPQRLNQLKSNKDR-DTKIFYSITGPGADSPPEGVAVE 161
Db 55 QRTKSSWVWQFVLEEYTGTDPLVYGLHSDMDRGDSIKYILSGEGAGI-----VFTID 110
QY 162 KETGWMLLNKPDLREBIAKVELFGHAVS-ENGASVEDPMNISIIVTDQNDHKPKFTQDTF 220
Db 111 DTTGDIHAIQRLDREERAQYTLRAQALDRRTGRPMPESEFEIKIQDINDNEPKFLDGPY 170
QY 221 RGSVLGVLPGTSVMQVATDDEDAIYTYGVVAVSIHSQEPKDPHDLMTFTHRSCTGIS 280
Db 171 VATVEMSPVGTSVIQVTATDADDDPTYGNSARVYVSILOQP-----YFSDSKTGVIR 224
QY 281 VISSGLDREKVPYTYLTIOATDMGDG--GSTTAVAVVEILDANDNAPMFPDPOKYEAHVP 338
Db 225 TALMNDREAKEVEYEVIIQAKDMGGQLGGLAGTTTNNITLSDVNDNPPRFPQKHQMSVL 284
QY 339 ENA-VGHEVQRLTVDLDAPNSPAMRATYILMGDDGDHFTTTHPESNQILTTKGLD 397
Db 285 ESAPISSTVGRVFAKDLDEGINAEMK--YTIVDGADAFDIDSTDPNFQVGIITVKKPLS 342
QY 398 FEAQOHTLYVEVTN---EAPFVLKLP-TSTATIVVHVEDVNEAPVFPVPSKVVEQEGI 453
Db 343 FESKSYTLKVGEGANPHLEMRFLNLGFPQDPTTVHLSVEDVDEPPVFEFGFYVEVEPDV 402
QY 454 PTGEPYCVYTAEDPKENQKISYRILR--DPAGWLAMPDPSGQVTAAGTLDREDEQFVRN 511
Db 403 AIGTTTIQIISAKDPDVTNNISIRYSDRSSDPGRFFVYDITTGALMTARPLDREPSW--- 459
QY 512 NIYEVWVLANMGSPPTTGTGTLLTLIDVNDHGP-VPEPRQITICNQSPVRHL---NI 567
Db 460 --HNITVLAEMNNPNSQVGSVPVTIKVLDVNDNAPEFPRFYFAVCENAKAGLIQTVSA 517
QY 568 TDKD 571
Db 518 VDQD 521

RESULT 8

US-11-108-172-1081
; Sequence 1081, Application US/111108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11/108,172
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 10/025,380
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 09/922,217
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/833,263
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 09/649,811
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 09/609,448

```

; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/575,251
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/519,444
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/504,629
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: US 09/480,321
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: US 09/476,296
; PRIOR FILING DATE: 1999-12-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1130
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1081
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-172-1081

Query Match          12.1%; Score 526.5; DB 7; Length 832;
Best Local Similarity 26.7%; Pred. No. 2.2e-32;
Matches 190; Conservative 109; Mismatches 308; Indels 105; Gaps 27;

Qy      59  PGQEPALF--STDNDDFTVRNGETVQE-----DWVVA-----PISVPENKGPFPQR 127
Db      141 PG-KPFLYVNVATDLDPATNGQLYYQIVQLPMINNMVYFQINNKTGAISLTREGSOEL 199

Qy      91  NPLKIFPSKRLRHKR-----DWVVA-----PISVPENKGPFPQR 127
Db      200 NPAK-NPSYNLVISVKDMGGQSENFSFDTTSVDIIVTENIWKAPKPVEMVNSTDPHPK 258

Qy      128 LNQLKSNKDRDTKIFYSTIGTGADSPPEGVFAVEKETGWLILNKPDLREEIAKTELFGHA 187
Db      259 ITQVRWN---DPGAQYSLV--DKEKLPFRFPESIDQE--GDIYVTPQLDREKDAVYFAYA 312

Qy      188 VSENGASVEDPMNISIIVTDQNDHKPKFTQDTFRGSLVGLVPGTSVMQVTTATDEDDAIY 247
Db      313 KDEVGKPLSYDLETHVKVKDINDNPCTPSPVTVFEQENRNLNSGTGLTAHDDRE-N 371

Qy      248 TYNGVWYVSIHQPKOPHDLMTHTHSTGTSTISVSSGLDRKREKPEYVTLTQATDMDGDG 307
Db      372 TANGFLNRYIVEQTPKPLPMDGLFLIQTVAGMLQAKQSLKKQDTPQYNLTIEVSDK--- 428

Qy      308 STTTAVAVVEILDANDNAMPDPQKYE--AHVPENAVGHEVQRLTVTDLDPNAPSAPKAT 365
Db      429 FKTLCFVQINVIDINDQIPIFEKSDYGNLTLAEDTNI GSTLTITQATDADEPFTGSSKIL 488

Qy      366 YLIMGDDGDHFTTTHPESNOGILTTKGLDFAKQHTLYVEVTNEAPPVLKL--PT 422
Db      489 YHIKGSSEGLGVDTDHTHTGVIYIKPLDFETAASVNIIVFKAENPEPLVFGVKYNAS 548

Qy      423 STATIVVHVEDVNEAPVFPSPKVEVQEGIPTGEPVCVYTAEDPDKENKISYRIILRDP 482
Db      549 SFAKFTLIVTDVNEAPQSQHVFQAKVSEDVAIGTKVGNVTAKDP--EGLDISYSLRGDT 606

Qy      483 AGWLAMPDPSQVAVGTLDREDSQFVRNNIYEVWVLAMDNQSPPTGTGTLTLITLIDVN 542
Db      607 RGLKIDHVTGEIFSVAPLDRE-----AGSPYRQVQVATEVGGSSLSVSEFFHILMDVN 661

Qy      543 DHGP--VPEPQITCN--QSPVRHLNITDKD---LSPH--TSPFQAQLTDDSDIYWT 592
Db      662 DNPRLAKDYTGLEFCHPLSAPGSLIFATDDQHLFRGPHFTFSLGSGSLQNDWEV--- 718

Qy      593 AEVNEBEGTVVLSSKKF--LKQDQYDVHLSDHGNK--EQLTVIRATVCDCHGHE--TCP 648
Db      719 SKIN--GTHARLSRTRHTDFEERAYVWLIRINDGGRPPLEGIVSLPVTFCSC---VEGSCF 773

Qy      649 GPWKGGFTLPVLGAVLALLF--LLLVLLLLLVKKRKKIKEPILLLPEDDTRDNY 598
Db      774 RPAGHQTGPTVGMAVGILLTLLVIGIILAVFIRIK-----KDKGKNV 819
```

```

RESULT 9
US-11-108-172-1087
; Sequence 1087, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11/108,172
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 10/025,380
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 09/922,217
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/833,263
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 09/649,811
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 09/609,448
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/575,251
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/519,444
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/504,629
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: US 09/480,321
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: US 09/476,296
; PRIOR FILING DATE: 1999-12-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1130
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1087
; LENGTH: 958
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-172-1087

Query Match          12.1%; Score 526.5; DB 7; Length 958;
Best Local Similarity 26.7%; Pred. No. 2.7e-32;
Matches 190; Conservative 109; Mismatches 308; Indels 105; Gaps 27;

Qy      59  PGQEPALF--STDNDDFTVRNGETVQE-----DWVVA-----PISVPENKGPFPQR 127
Db      267 PG-KPFLYVNVATDLDPATNGQLYYQIVQLPMINNMVYFQINNKTGAISLTREGSOEL 325

Qy      91  NPLKIFPSKRLRHKR-----DWVVA-----PISVPENKGPFPQR 127
Db      326 NPAK-NPSYNLVISVKDMGGQSENFSFDTTSVDIIVTENIWKAPKPVEMVNSTDPHPK 384

Qy      128 LNQLKSNKDRDTKIFYSTIGTGADSPPEGVFAVEKETGWLILNKPDLREEIAKTELFGHA 187
Db      385 ITQVRWN---DPGAQYSLV--DKEKLPFRFPESIDQE--GDIYVTPQLDREKDAVYFAYA 438

Qy      188 VSENGASVEDPMNISIIVTDQNDHKPKFTQDTFRGSLVGLVPGTSVMQVTTATDEDDAIY 247
```


Sequence 98, Application US/11102240
Publication No. US20050260647A1
GENERAL INFORMATION:
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
TITLE OF INVENTION: ESOPHAGEAL TUMOR
FILE REFERENCE: P3230KIC106C
CURRENT APPLICATION NUMBER: US/11/102,240
CURRENT FILING DATE: 2005-04-08
PRIOR APPLICATION NUMBER: 10/063662
PRIOR FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 10/006867
PRIOR FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 199-12-09
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 98
LENGTH: 807
TYPE: PRT
ORGANISM: Homo Sapien
US-11-102-240-98

Query Match 8.7%; Score 381.5; DB 7; Length 807;
Best Local Similarity 20.9%; Pred. No. 2e-21;
Matches 170; Conservative 99; Mismatches 251; Indels 293; Gaps 25;

Qy 114 ISVPENKGPORLNOLKSKND-RDTKIFVSIITGPGADS--PPEGVFAVEKETGWMLLN 170
Db 27 VEVPENTGNPFLYTLKLPREGAQIVLUS-----GDSKATEGPFAMDPDSGFLVLT 81

Qy 171 KPLDREBIAYELFGHAVSENGAGVEDPMNIIIVTDQNDHKPKFTQDTPRGSLVGLVLP 230
Db 82 RALDREAEQVQLQVLTLEMQDGHVLMGPQVLVHVHKENDQVPHFSQAIYRARLSRGRTP 141

Qy 231 GTSVMQVATDEDDAIYTYNGVAYSISHPKDPKPHLMFTI----- 272
Db 142 GIPELFLEASDRDEP-GTANSDLRPHILSOAPQSPDMFOLEPRLGALALSPKGSSTLD 200

Qy 273 -----HSTGTISV----- 281
Db 201 HALERTYQLLVQVKMDGQASGHQATATVEVSIITWVSLEPIHLAENLKVLYPHHMAQ 260

Qy 282 -----ISSGLDREKVPYETLTITQATMDGDSGTTTA 312
Db 261 VHWSGDVHYHLESHPPGPFVNAEGNLVYTRDLREAQAEYLLQVRAQNSHGSDYAAPL 320

Qy 313 VAVVEILDANDNAPMFPQKYEAHVE-NAVGHVEVQLTVTTDLDPNSPAWRATYLLIM-- 369
Db 321 ELHLVMDNDNVPICPDRPTVSIPELSPGTEVTRLSAEDADAPGSPNSHVYQLLSP 380

Qy 370 -----CG----- 371
Db 381 EPEDEVEGRAFOVDPTSGSVTLGVPLRAGQNILLVLMDLAAGGFGSSTCEVEAVT 440

Qy 372 DDGDH-----FT 378
Db 441 DINHAPFITTSQIGTSLPDEVBPGTLVAMLTADADLEPAFLMDFAIERGDTGTFG 500

Qy 379 ITTHPESNGGILTRKGLDFRKNQHTLYVEVTNEAPVVKI-----PTSTATIVVHVE 432
Db 501 LDWEPDSGHVRLRLCKNLSEYAAPSHVWVWVQS-----VAKLVGPGPGCATATVTVLVE 556

Qy 433 DVNEAPVFPSPKVE--VQSGIPTGEPV--CVVTAEDPPKENQKISVRIILRDPAGWLM 488
Db 557 RV-----MPPPKLDQSEYASVAPAGSFLTLTIQPSDPISTRFLPSLVNDSEGLICI 610

Qy 489 DPDSQGVAVGTLDREDBQFVRNNIYEVWVLMADNGSPPTTGTGTGLTLTLTLIDVNDHGVPV 548

Db 611 EKFSGEVHTAQSLQGAQP-----GDTYTVLVVEAQDTA-----LTL-----APVP 649

Qy 549 E-----PRQITICNQSPVRHLNIT-----DKDLSPHTSFPQALTDSDIY--WTAEVNE 597

Db 650 SQYLCTPRQ-----DHGLIVSGSPKDPDLSAGHGPYSFTLGNPTVQORDWRLOTLN 700

Qy 598 EGDTVVLSLKKFLKQDQTYDVHLSLSDHGKNEQLTVIRATVCDCH--GHVETCFGPMKG-- 653

Db 701 GSHAYLTALHWHVEPREHIIPVVVSHNAQMQL-LVRVIVCRNVEGQCQMKVGRMKGMP 759

Qy 654 -----GFILPVLGAVALLFLLLVLLVLRKK 680

Db 760 TKLSAVGILVGTILVAIGIFILIFLIFTHWTMSRKK 792

RESULT 12
US-10-131-826A-394
Sequence 394, Application US/10131826A
Publication No. US20050245730A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: F3330RIC128
CURRENT APPLICATION NUMBER: US/10/131,826A
CURRENT FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 394
LENGTH: 1184
TYPE: PRT
ORGANISM: Homo Sapien
US-10-131-826A-394

Query Match 7.6%; Score 333.5; DB 6; Length 1184;
Best Local Similarity 21.1%; Pred. No. 1.5e-17;

Matches 191; Conservative 139; Mismatches 378; Indels 199; Gaps 36;	
QY	6 GPLASLLLVQV-----LQAASEPCCRAVREAEVTLAAGAEQEPGQALGVFMG 57
Db	13 GPGGYLFLGDCQEVTTLVKYQVSEEVPSGTIGKLSQEL---GREERRQA-GAAFOV 68
QY	58 CPQEPALFSTDNDDFTVRNGETVQERSLSKERPLKIFPSKILRRHKRDWVAISVP 117
Db	69 LQLPQALPIQVDSSEGLSTGRELDREQLCRQWDPCLV--SPDVLATGDIALHVEIQVL 126
QY	118 E-NGKGP-FPQRLNQLKSNKDRDTKIFYSI-----TGPCA---DSPPEGVFAV--- 160
Db	127 DINDHQPRFPKQGELEISSASLRTPIDLRALDDPTGENTLHTLSPSEHFALDVIV 186
QY	161 ---EKETGMLLNKPLDREBIAKYLFGHVAENGASVEDPMNISIIIVTDQNDHKPKFTQ 217
Db	187 GPDTKHAELIVVKELDREIHSFFDLVLTAYDNGNPKSGTSLVKVNLDSNDNSPAFAE 246
QY	218 DTRGSLVLEGVLGTSMQVATDEDDAIYTYNGVAVSIHSQEPKDPDPLMTIHRSTG 277
Db	247 SSLALEIQEDAAPGTLILKLTATDPDQG---PNGEVEFFLCKMPPEVLD-TFSIDAKTG 302
QY	278 TISVSSGLDREKVPYTLTIQATDMGDGSGTTTAVAVVEILDANDNAPMF-----DPQ 331
Db	303 QV-ILRPLDYENKPAVEVDVQARDLGPNIPIAHCKVLKVLVDVNDNIPISIHVTWASQPS 361
QY	332 KYEAVHPENAVGHEVORLVTDLDPNS---PAWRATYILMGDDGDGHFTTITTHPSNQ 388
Db	362 LVSEALPKDSF---IALWADDLDSGHGLVHCLSQEL-----GHFRL-KRTNGNTY 410
QY	389 ILTRKGLDPEAKNQHTLYVEVNEAPFVLKLTSTATIVVHVEDNE-APVPVPSKV 447
Db	411 MLLTNATLDREQWPKYTLTLAQDQG---LQLSAKKQLSIQISDINDNAPVEKRYEV 467
QY	448 EVOEGITGPVPCVYTAEDPDKE-NOKISYRILRDPAGWL-AMDPDGGVTAAGTLDRED 505
Db	468 STRENNLPSLHITIKAHDADLNGKVSRIQDSPVAHLVAIDSNTEVTAORSUNYEE 527
QY	506 EQFVRNNIYEMVWLAMNGSPPTGTGTLTLLIDVNDHGP-VPEP-----RQITICNQ 558
Db	528 MAG-----PEFQVIAEDSGQPLASSVSVWVSLLDANDNAPEVQPVLSGKASLSVLVN 582
QY	559 SPVRHVL-----NITDKD----- 571
Db	583 ASTGHLVPIETPNGLGPAGCTTPPLATHSSRPFLTTIIVARDADSGANGEPLYSIRNGN 642
QY	572 -----LSPHTSPFQQLTDDSDIYWTAEVNEEGDVTVLKFLKQDVTVDHLSLSDHG 625
Db	643 EAHFLINPHTGOLFVNVTNASSL-----IGSE-----WELIIVVEDQG 681
QY	626 N-----KEQLTVIRATVCDCHGHVETCPGPKWGGFILPVLGAVLALLF-LLLVLLLVK 679
Db	682 SPPLQTRALLRVMFVTSVDHLRDSARKPGALSMLTVICLAVLLGIFGLILALFMSICR 741
QY	680 KRKIKPLLLPEDDTRNDVPYGEGBEDQDYDTQLHRGLEARPEVVLRNDVAPTII 739
Db	742 TEK-----KDNRAYNCREA-----ESTYRQPKPKQKHQK---ADIH 776
QY	740 PTWYRPRANPDEIG-----NFIENLKAANTDP-----TAPPYDTLLVPDYGSGS 787
Db	777 LVPVLRGAGEPCEVGQSHKVDKEAMMEAGWDPCLOAPHLPTTLTYRTL-----RNQGN 831
QY	788 DAASLSS 794
Db	832 QGAPAES 838

RESULT 13

US-10-131-826A-130

; Sequence 130, Application US/10131826A

; Publication No. US20050245730A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen	
; APPLICANT: DeForge, Laura	
; APPLICANT: Desnoyers, Luc	
; APPLICANT: Filvaroff, Ellen	
; APPLICANT: Gao, Wei-Qiang	
; APPLICANT: Gerritsen, Mary E.	
; APPLICANT: Goddard, Audrey	
; APPLICANT: Godowski, Paul J.	
; APPLICANT: Gurney, Austin L.	
; APPLICANT: Sherwood, Steven	
; APPLICANT: Smith, Victoria	
; APPLICANT: Stewart, Timothy A.	
; APPLICANT: Tumas, Daniel	
; APPLICANT: Watanabe, Colin K	
; APPLICANT: Wood, William	
; APPLICANT: Zhang, Zemin	
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC	
; FILE OF INVENTION: ACIDS ENCODING THE SAME	
; FILE REFERENCE: P3330R1C128	
; CURRENT APPLICATION NUMBER: US/10/131,826A	
; CURRENT FILING DATE: 2002-04-24	
; PRIOR APPLICATION NUMBER: 60/049911	
; PRIOR FILING DATE: 1997-06-18	
; PRIOR APPLICATION NUMBER: 60/056974	
; PRIOR FILING DATE: 1997-08-26	
; PRIOR APPLICATION NUMBER: 60/059113	
; PRIOR FILING DATE: 1997-09-17	
; PRIOR APPLICATION NUMBER: 60/059115	
; PRIOR FILING DATE: 1997-09-17	
; PRIOR APPLICATION NUMBER: 60/059184	
; PRIOR FILING DATE: 1997-09-17	
; PRIOR APPLICATION NUMBER: 60/059117	
; PRIOR FILING DATE: 1997-09-17	
; PRIOR APPLICATION NUMBER: 60/059122	
; PRIOR FILING DATE: 1997-09-17	
; PRIOR APPLICATION NUMBER: 60/059184	
; PRIOR FILING DATE: 1997-09-17	
; PRIOR APPLICATION NUMBER: 60/059263	
; PRIOR FILING DATE: 1997-09-18	
; PRIOR APPLICATION NUMBER: 60/059352	
; PRIOR FILING DATE: 1997-09-19	
; PRIOR APPLICATION NUMBER: 60/059588	
; PRIOR FILING DATE: 1997-09-19	
; Remaining Prior Application data removed - See File wrapper or PALM.	
; NUMBER OF SEQ ID NOS: 550	
; SEQ ID NO 130	
; TYPE: PRT	
; LENGTH: 530	
; ORGANISM: Homo Sapien	
; US-10-131-826A-130	

Query Match 6.4%; Score 280; DB 6; Length 530;	
Best Local Similarity 27.2%; Pred. No. 5.3e-14;	
Matches 126; Conservative 65; Mismatches 200; Indels 72; Gaps 23;	
QY	208 QNDHRPKFTQDTRGSLV---EGVLPTGSMQVATDEDDAIYTYNGVAVSIHSQEPKDP 265
Db	24 QVNRLPFFTHFFDFTYLLISEDTPGVSSVTQLLAQDMD-----NDPLVFGVSGEEA--- 74
QY	266 HDLMFTIHRSTGTISVISSGLDREKVPYTLTIQATDMGDGSGTTTAVAVVEILDANDNA 325
Db	75 -SRFAVEPDGTUV-WLRQPLDRETKSEFTVFSVDHQG---VITRKNIQVGDVNDNA 129
QY	326 PMFDPQKYBAHPENA-VGHEVQRLTVTDLDPNSPAWRATYILMGCD-----DGDHF 377
Db	130 PTFHNPQSVRIPEPTPVGTPIFVNATDPD-----LGAGGSVLVSFPSPSQFF 178
QY	378 TITTHPESNOGLITRKGDLDFEAKQOHTLYVEVNEAPFVLKLTPTST-----ATIVVHVED 433
Db	179 AI-----DSARGIVTVIRELDYETQAYQLTVNATDQDK---TRPLSTLANLAIITVDQD 231
QY	434 VNEAPVFPVPSKVVEQEGIPTEGPEVCVVTAEADPKKENOK-ISRILRDPAGWL-AMDPD 491
Db	232 MD--PIFINLPYSTNIYHSPPGTTVRIITAITDQDKRPRGIGYTIIVSGNTNSIFALDIYI 289

Db 1252 RTVTAFVEKEG-----MLEEAEPLRASDRNIMCEDDCHDTYYYSIVGNGSGHFTV--D 1305
Qy 383 PESNOGILTRKGLDFEAKNQHTLYVEVTNEAPFVLKLPSTSTATIVVHVHVEDVNEAPVFPV 442
Db 1306 PRTN--VLSLVKPLDRSEQETHLIIGASDTNPAAVLQASTLTVTVVNREANPRPVF-- 1361
Qy 443 PSKVVEQEGIPGEPV--CVYTAEDDDKENQKISYRILRDPAGWLAMDPSGQVTAVGT 500
Db 1362 --QRALYTAGISAGDFIERNLTLVATHSEDLPTVTLIOE-----SMEADP-----T 1407
Qy 501 LDR-EDeqfVRN-----NIVVMVLAMDNQSPPTGTGTLTLLTLIDVN 542
Db 1408 LEAVQSAFLNPEVGTGVLNFOPTASMHGMEFEVKAADSRT-ETARTEVKVILISDRN 1466
Qy 543 D-----HGPVPE--PROITI-----CN-----QSPVRHVLNITDKLSPHTS 577
Db 1467 RVFFTFNNPLPEVTPQEDFIAETFTAFFGTCNIDQSWASDPVTGATKDDQTEVRAHP- 1525
Qy 578 PFOAQLTDDSDIYWTAEVNEE---GDTVVLSLKKFLKQDTYDVHLSLDHGNKEQLTVIR 634
Db 1526 -----IRDDLFPV--PAEEIEQLRGNPTLVNSIQRALEEQ---NLQLAD----- 1563
Qy 635 ATVCDCHGHVETCPGPKG-----FILPVLGAVLALLFLLLVLLLVKXK---RKIK 684
Db 1564 -----LFTGETPILGGDAQARALYALAAVAAALIVVLLIVFVRTETLNRL- 1613
Qy 685 EPLLPEDDTRDN-----VFYEGEGGGE-----EDQDYDITQLH 719
Db 1614 QALSMTKYSSQDSGLNRVGLAAPTGNKHAVEGSPINWNETLKAPDFDALSEQSYDSDLI- 1672
Qy 720 RGLERPEVVLNRNDVAPTIPTMYRPRAN-----PDEIGNFIENLKAANTDP 769
Db 1673 -GIEDLPQ--FRNDYFP-----PEGSSMRGVVNEHVPESIANH--NNNFGFNSTP 1718
Qy 770 TAPPY 774
Db 1719 FSPEF 1723

Search completed: December 12, 2005, 09:20:10
Job time : 15 secs

THIS PAGE BLANK (USPTO)